De novo mutations of *MSL3* cause a X-linked syndrome marked by impaired histone H4 lysine 16 acetylation

M. Felicia Basilicata[1], Ange-Line Bruel[2*], Giuseppe Semplicio[1*], Claudia Isabelle Keller Valsecchi[1*], Tuğçe Aktaş[1*], Yannis Duffourd[2], Tobias Rumpf[1], Jenny Morton[3], Iben Bache[4,5], Witold G. Szymanski[1], Christian Gilissen[6], Olivier Vanakker[7], Katrin Õunap[8], Gerhard Mittler[1], Ineke van der Burgt[6], Salima El Chehadeh-Djebbar[2,9], Megan T Cho[10], Rolph Pfundt[6], Tiong Yang Tan[11], Maria Kirchhoff[4], Björn Menten[7], Sarah Vergult[7], Kristin Lindstrom[12], André Reis[13], Diana S. Johnson[14], Alan Fryer[15], Victoria McKay[15], DDD Study[16], Richard B. Fisher[17], Christel Thauvin-Robinet[2], David Francis[18], Tony Roscioli[19,20,21], Sander Pajusalu[8], Kelly Radtke[22], Jaya Ganesh[23], Han G. Brunner[6,24] , Meredith Wilson[25], Laurence Faivre[2], Vera M. Kalscheuer[26], Julien Thevenon[2, 27#] and Asifa Akhtar[1#]

[1]Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany
[2]Inserm UMR 1231 GAD, Genetics of Developmental disorders, Université de Bourgogne-Franche Comté, FHU TRANSLAD, Dijon, France + Centre de Référence Maladies Rares "Anomalies du Développement et syndromes malformatifs", FHU-TRANSLAD, CHU Dijon Bourgogne, France
[3]West Midlands Regional Clinical Genetics Service and Birmingham Health Partners Birmingham Women's Hospital NHS Foundation Trust, UK
[4]Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
[5]Wilhelm Johannsen Centre for Functional Genome Research, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark
[6]Department of Human Genetics, Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, 6500 GA, The Netherlands.
[7]Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium
[8]Department of Clinical Genetics, United Laboratories, Tartu University Hospital and Institute of Clinical Medicine, University of Tartu, Tartu 51014, Estonia
[9]Service de Génétique Médicale, Hôpital de Hautepierre, Strasbourg, France
[10]GeneDx, Gaithersburg, Maryland 20877, USA.
[11]Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Royal Children's Hospital, University of Melbourne Department of Paediatrics, Flemington Road, Parkville VIC 3052 Australia
[12]Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, USA
[13]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany.
[14]Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, U.K
[15]Department of Clinical Genetics, Liverpool Women's NHS Foundation Trust, Liverpool, L8 7SS, UK
[16]DDD Study, The Wellcome Sanger Institute, Hinxton, Cambridge, UK
[17]Northern Genetics Service, Teesside Genetics Unit, The James Cook University Hospital, Middlesbrough TS4 3BW, UK

[18]Cytogenetic laboratory, Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Royal Children's Hospital, Flemington Road, Parkville VIC 3052 Australia
[19]Neuroscience Research Australia, Margarete Ainsworth Building, Barker Street Randwick Sydney NSW 2031 Australia
[20]Prince of Wales Clinical School, University of New South Wales, NSW 2031 Australia
[21]Department of Medical Genetics, Sydney Children's Hospital, High Street, Randwick NSW 2031, Australia
[22]Department of Clinical Genomics, Ambry Genetics, Aliso Viejo, California, USA
[23]Division of Genetics, Cooper University Hospital & Cooper Medical School at Rowan University, Camden, NJ 08103 USA
[24] Department of Clinical Genetics and School for Oncology and Developmental Biology (GROW), Maastricht University Medical Center, 6202 AZ, Maastricht, The Netherlands
[25]Department of Clinical Genetics, Children's Hospital at Westmead, Disciplines of Genetic Medicine and Child and Adolescent Health, University of Sydney, Sydney, Australia
[26]Research Group Development and Disease, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany
[27]Centre de génétique, CHU Grenoble; CNRS UMR 5309, INSERM, U1209, Université Grenoble Alpes, Institute of Advanced Biosciences, Grenoble, France

* These authors contributed equally to this study

Co-Corresponding authors
jthevenon@chu-grenoble.fr
https://orcid.org/0000-0001-9271-3961
Phone: (33) 0476767285
Fax: (33) 0476767286

akhtar@ie-freiburg.mpg.de
https://orcid.org/0000-0001-8973-6193
Phone: (49) 07615108565
Fax: (49) 07615108566

Abstract

The etiological spectrum of ultra-rare developmental disorders remains to be fully defined. Chromatin-regulatory mechanisms maintain cellular identity and function, where misregulation may lead to developmental defects. Here, we report pathogenic variations in *MSL3*, which encodes a member of the chromatin-associated MSL-complex responsible for bulk Histone 4 Lysine 16 acetylation (H4K16ac) in flies and mammals. These variants cause a novel X-linked syndrome affecting both sexes. Clinical features of the syndrome include global developmental delay, progressive gait disturbance and recognizable facial dysmorphism. MSL3 mutations affect MSL-complex assembly and activity, accompanied by a pronounced loss of H4K16ac levels *in vivo*. Patient-derived cells display global transcriptome alterations of pathways involved in morphogenesis and cell migration. Finally, we use histone-deacetylase inhibitors to rebalance acetylation levels alleviating some of the molecular and cellular phenotypes of patient cells. Taken together, we characterize a novel syndrome allowing us to decipher the developmental importance of MSL3 in humans.

Introduction

Recent advances in whole-exome sequencing have led to discovery of numerous causative genes for human diseases. However, gathering a large number of patients represents a significant bottleneck for gene identification and appropriate molecular analysis of rare developmental diseases. Mutations of epigenetic regulators were found to cause neurodevelopmental disorders through mechanisms potentially altering the chromatin landscape[1]. To date, over forty Mendelian disorders have been identified, which are caused by mutations of the epigenetic machinery and are associated with neurological dysfunction, with 80% of them being autosomally encoded[2–4]. Fluctuations in the histone acetylation balance, due to the interplay between acetylase and deacetylase functions during development, contributes to the final shape and performance of the adult brain and the whole organism[5–8]. Whilst several modifications within the histone core were found to alter chromatin architecture[9], Histone 4 Lysine 16 acetylation (H4K16ac) appears uncommon

amongst the tail modifications as it directly affects higher order chromatin structure[10]. The Male Specific Lethal (MSL) complex-associated MOF/KAT8 histone acetyl-transferase is responsible for bulk H4K16ac in flies and mammals[11]. Given its crucial importance, we know surprisingly little about the role of MSL-mediated H4K16ac during development and whether its misregulation is associated with any disease in humans.

Results

Mutations in *MSL3* cause a novel neurodevelopmental disorder

Through the international Matchmaker Exchange Initiative[12], we have identified *de novo* variants in *MSL3* (Male Specific Lethal homolog 3 Homolog, previously known as *MSL3L1*, Refseq gene NG_012564.1) in fifteen unrelated families from Europe, the US and Australia (Figures 1a, 1b, Supplementary Figure 1a). In the matchmaking cohort we identified eleven truncating variants (p1-3, p5-11, p13), one exon deletion (p12), one missense variant (p4), two gene-deletions (p14-15) and one chromosomal inversion (p16) (Figure 1c, Table 1). Patients 1 to 9 were gathered from a cohort of approximately 10,000 individuals with intellectual disability, demonstrating a significant enrichment for *MSL3 de novo* truncating variants in this cohort compared to random occurrence (uncorrected p-value = 8.98e-16, p_{FDR} = 1.64e-11*)* (Supplementary Figure 1b-c, Table 1). The enrichment for truncating variants in this cohort of patients is consistent with the data from the ExAC database, suggesting that *MSL3* is extremely intolerant to loss-of-function variations (pLI=0,95)[13]. For patients 5 and 6 (monozygotic twins), deep sequencing of leucocyte parental DNA did not detect an alternative allele. Patient 16 inherited the X-inversion causing a putative in-frame transcript between *MSL3* and *GAB3* from his mildly affected mother (Table 1 and Supplementary Figure 1d).

MSL3 is known for its role during X chromosome dosage compensation via H4K16ac in *Drosophila*. Its absence affects male flies and manifests in lethality in the last larval stage of development[11,14,15]. The human *MSL3* gene resides on the X chromosome and its mutations presented here affect both hemizygous male and heterozygous female individuals (Figure 1a, Supplementary Table 1). Common features across the 16 patients include hypotonia, feeding difficulties in early infancy

and global delay in the acquisition of major milestones including walking and speech. Intellectual disability was also scored. The phenotypic spectrum is more homogeneous across patients of both sexes with single nucleotide variants (SNVs) compared to patients with chromosomal rearrangements. The patients with SNVs display progressive spasticity or an ataxic gait (9/13), resulting in inability to walk (5/13). The facial morphology is strongly overlapping across patients with telecanthus (8/13), epicanthal folds (9/13), down-slanting palpebral fissures (9/13), downturned corners of the mouth (8/13) and dysplastic ears (11/13) (Figure 1b, Supplementary Table 1, also see the detailed clinical descriptions in Supplementary Note 1). Altogether, these clinical observations delineate a recognizable *MSL3* syndrome. In line with these phenotypes, *MSL3* is ubiquitously expressed in the developing human brain until 17-20 weeks post conception. Other MSL complex members display a similar expression pattern, though their RNA levels seem slightly higher (Supplementary Figure 1e[16]).

Most of the pathogenic *MSL3* variants affect the C-terminal MRG domain (Figure 1d, Supplementary Figure 1f), which is responsible for MSL complex formation while enhancing the activity of MOF towards H4K16ac (Figure 1e,[17,18]).

Using *in silico* modeling of the crystal structure of the human MSL3-MSL1 complex[17], we predicted that the truncations present in the patients impair the integration of mutant MSL3 into the MSL complex (Figure 1f, g). To test this, we transiently transfected expression vectors encoding wild-type (full-length) and three different MSL3 mutants, representing patients 1, 2 and 3 (p1, p2, p3) into Flp-In T-REx 293 cells. We observed that in contrast to wild-type MSL3, which showed robust interaction with MSL1 and MOF, all three mutants lost the interaction with MOF and to a lesser extent with MSL1 (Figure 1h, lanes 6-9). Two of the mutations also decreased MSL3 protein stability (Figure 1h, lanes 2 and 3). Taken together, we find that the mutations present in the patients with *MSL3* syndrome lead to severely compromised MSL complex integrity.

Bulk reduction of H4K16 acetylation in *MSL3* patient cells

To characterize the molecular consequences of *MSL3* mutations in patients, we obtained skin biopsies from two female patients harboring frameshift mutations (p1 and p2), one female patient with a locus deletion (p14) and one control tissue (ctrl) from p1's healthy mother. In addition, we analysed two more skin biopsies from

healthy donors as controls (see Methods). These samples were used to perform histological analyses and establish primary dermal fibroblast cell culture models for all molecular analyses presented below.

Hematoxylin and Eosin staining confirmed that the overall tissue architecture was preserved in both control and patient specimens (Supplementary Figure 2a). The MSL complex is responsible for bulk H4K16ac in flies[19,20] and mammalian cell culture models[14,15,21]. Based on this, we performed immunostaining of H4K16ac in patient samples compared to skin punches of controls. While the H4K16ac staining had homogeneous intensity in dermal and epidermal layers of all three control samples, *MSL3* patient cells showed a heterogeneous pattern with a marked and variable reduction in H4K16ac (Figure 2a, asterisks). This striking reduction of H4K16ac could result from decreased levels of MOF in *MSL3* patient cells. However, MOF immunostaining displayed a homogenous distribution within all cells of the tissue section (Figure 2b). This indicates that the loss of H4K16ac is not caused by the reduction in MOF levels *per se* but rather due to a misregulation of its enzymatic activity.

We confirmed these observations in primary fibroblast cells established from skin punches, where cells lacking functional MSL3 also displayed a severe decrease in H4K16ac (Figure 2c, d), but not other histone modifications (e.g. H3K27me3, Figure 2e and Supplementary Figure 2b). Remarkably, quantification of these stainings showed a global reduction of H4K16ac in these heterozygous $X^{MSL3\text{-WT}}/X^{MSL3\text{-MUT}}$ female patient samples (Figure 2d). Such a staining pattern is unusual for an X-linked gene in females, where conventionally, heterozygous mutations would be expected to result in a 50:50 population of wild-type and mutant cells. Targeted sequencing of *MSL3* cDNA for patient 1 and 2 demonstrated the mild co-expression of mutated and wild-type allele at the RNA level (Supplementary Figure 2d). While the nature of *MSL3* mutation in p14 (whole gene deletion) precludes expression of any mutated MSL3 protein, we note that p1/2 proteins were highly unstable (Figure 1h), making it unlikely that the *MSL3* syndrome is generally caused by a dominant negative effect. However, human *MSL3* was recently found to be a facultative escapee[22], which is stochastically and variably expressed from both female X chromosomes. We therefore propose that *MSL3* haploinsufficiency might cause the variable but consistent reduction of H4K16ac in female patients.

Since there could be a potential crosstalk between H4K16ac and other residues on the histone tails[23] and given that MSL3 has been linked to H3K36me3 and H4K20me1[24,25], we validated the above findings by bulk Western blot analyses and quantitative mass spectrometry (LC-MS). In line with the immunofluorescence, we exclusively scored a reduction of MSL3 and H4K16ac levels, while MOF and other modifications remained unaffected in western blots (Figure 2f, Supplementary Figure 2c, Supplementary Figure 5). Affirming these results, when comparing the three patients to two different controls by comprehensive histone LC-MS analysis, the modification that was by far the most significantly affected, both by p-value and fold-change, was H4K16ac (Figure 2g, Supplementary Figure 2e, f and Supplementary Table 2). Hence, loss of MSL3 in patients appears to primarily affect H4K16ac.

H4K16ac has been previously linked to cell cycle regulation[21,26–28]. We noticed that patient-derived fibroblasts grew significantly slower than control cells (Supplementary Figure 2g), nevertheless no clear difference of any specific cell cycle phase could be scored (Supplementary Figure 2h). Cells did not show signs of G1/S arrest both by qPCR for senescence markers nor by X-Gal staining (Supplementary Figure 2i, j). Furthermore, no differences were scored for γH2AX signal, indicating no apparent increase in DNA damage (data not shown).

In order to further dissect the possible cause for the slow growth of female *MSL3* mutant cells, we next challenged them by removal of serum and tested their capacity to respond to changes in environmental conditions. Interestingly, upon performing these starvation experiments we observed that patient-derived fibroblasts showed a delay in exiting G0 phase (Supplementary Figure 2k). This suggests that *MSL3* mutant cells may fail to timely respond to growth factor signals, a property that is crucial during development.

Transcriptome deregulation in *MSL3* patient cells

Since MSL3 is part of a chromatin-modifying complex that is involved in transcription regulation[21,29,30], we next performed RNA-seq experiments on three patient-derived primary fibroblast cell lines. For this purpose, we analysed two different passages compared to a passage matched control (ctrl) in synchronized conditions. Differential expression analysis was performed using DESeq2 and genes with FDR correction <0.05 were considered differentially expressed (DE) (Figure 3a). As gene expression differences can reflect genetic background or environmental influences between

individuals[31], we decided to consider only transcripts that were consistently deregulated in all three patients. This revealed a set of 519 commonly misregulated genes (Figure 3a,b, Supplementary Figure 3a).

We first explored whether altered expression of these genes could be directly related to MSL-mediated H4K16ac. We noticed that the extent of misregulation (log2FC) was much more consistent between different patients in the downregulated subset compared with the upregulated group of genes (Figure 3a and c). This supports the idea that the latter arise from secondary responses within each individual, which may be more variable. We also found that the orthologues of the DE-genes in *MSL3* patients were significantly enriched for mMSL2 ChIP peaks in mouse cells[21]. In contrast, mKANSL3 peaks were significantly underrepresented, corroborating the hypothesis that the observed effects are caused by H4K16ac deposited by the MSL-, but not the KANSL-complex (Figure 3a, d). Therefore, we performed H4K16ac ChIP experiments in *MSL3* patient cells and focused on conserved mammalian MSL targets that are mMSL2-bound in mouse cells and very consistently downregulated in both *MSL3* patients as well as upon *MSL3* knock-down in male and female fibroblasts (Figure 3e). These results substantiated our conclusion, as H4K16ac was lost from *ZNF185* and *SPON2*, but not *H3F3B* (Figure 3e, Supplementary Figure 3b).

We next focused on analyzing the biological functions and pathways orchestrated by the DE-genes in *MSL3* syndrome. Intriguingly, we found differentially expressed Notch-, Hox- and Wnt family members, which are key morphogenetic regulators of development from *Drosophila* to mammals. We confirmed these findings by RT-qPCR, for which we included several additional control cell lines (ctrl2, ctrl3). Remarkably, patient cell lines showed very consistent behavior in the selected set of downregulated as well as unaffected control genes (Figure 3f, Supplementary Figure 3c, Supplementary Table 5). We could also confirm the reduction at the protein level for HTR7 and UNC5B in the patient skin specimens (Supplementary Figure 3d).

Pathway perturbation analysis revealed “Dermatological Diseases”, “Organismal Abnormalities”, “Gastrointestinal Disease” as well as “Neurological Disease” as top scoring disease signatures (Figure 3g). More specifically, the genes commonly affected upon *MSL3* mutation are overall involved in the activation of cellular movement, the organization of the cytoskeleton or the formation of cell protrusions, while mechanisms to reduce cell death are repressed (Figure 3h). These are all characteristic symptoms of the *MSL3* syndrome underscoring the relevance of our

RNA-seq dataset (see Supplementary Table 3 for DE- and clinically relevant genes based on *MSL3* patient phenotypes).

We were next interested in dissecting the primary gene expression perturbations caused by acute loss of *MSL3* versus secondary effects that may accumulate during development in the patients. Therefore, we performed *MSL3* knock-down experiments in control primary fibroblast cell lines as well as *MSL3* rescue experiments in patient cells to determine the extent to which we could phenocopy the gene expression changes detected by RNA-seq. We observed an acute downregulation of multiple target genes (e.g. *ZNF185*, *SPON2*, *NOTCH3*, *FAM157*) upon *MSL3* knock-down in both male and female cells (Supplementary Figure 3e). Genes found to be upregulated in *MSL3* patients typically did not respond (e.g. *PRICKLE*) or were sometimes downregulated (*WNT2*). When we did the inverse experiment by providing *MSL3* in patient cells, we noted a similar trend (Figure 4a). We could rescue several downregulated targets (*ZNF185*, *SPON2*, *FOSB*, *EGR2*) by re-expressing *MSL3*. Interestingly, these genes were also upregulated in ctrl cells, indicating that they might be more generally susceptible to alteration of MSL3 and H4K16ac levels. Taken together, our gene expression analysis suggests that lack of MSL complex results in acute downregulation of several biologically relevant genes including key developmental regulators. This misregulation may then gradually trigger secondary perturbances, collectively resulting in the global disease state of the *MSL3* syndrome.

HDACi-mediated alleviation of phenotypes in patient cells

Given the clear link between H4K16ac and *MSL3* mutation, we considered whether it would be possible to not only rescue the phenotypes of patient cells by reintroducing *MSL3* (Figure 4a) but also by re-adjusting acetylation levels. We therefore tested three different classes of histone deacetylase inhibitors (HDACi), which are currently approved for Clinical Trial Phase II or III: SAHA (vorinostat)[32,33], LBH-589 (panobinostat)[34], MGCD0103 (mocetinostat)[35], as well as Trichostatin A (TSA) as a positive control[36,37]. These inhibitors show strong effects on histone H4 acetylation and in particular H4K16ac in human transformed cells[37] (Supplementary Figure 4a). Hence, we referred to the aforementioned study for using HDACi to modulate H4K16ac in our primary untransformed cells.

Treatment of control and patient cells with HDACi resulted in bulk histone hyperacetylation and triggered upregulation of acetylation-sensitive genes (Supplementary Figure 4b, c, d). We first verified that the response to HDACi is effective in all mutant cells, irrespective of the MSL3/H4K16ac levels. We therefore performed immunostainings on HDACi-treated patient cells and observed an overall increase of H4K16ac at the single cell level from both MSL3 high and low populations (Figure 4b). Quantification of those stainings in patient-derived cells confirmed the rescue of H4K16ac, which now was comparable to healthy controls (Figure 4c). We next tested whether this has a positive influence on gene expression. Remarkably, treatment with HDACi specifically rescued the expression level of selected downregulated targets whilst control genes or upregulated targets remained largely unchanged (Supplementary Figure 4e and data not shown). In contrast to p1, p2 and p14 patient cells responded more selectively to LBH-589 (panobinostat), while the other two HDACi (SAHA and MGCD0103) showed only a modest rescue of gene expression defects (Supplementary Figure 4e). To characterize the HDACi response in *MSL3* deficient cells in a more comprehensive manner, we performed RNA-seq upon treatment with LBH-589. As expected, HDACi treatment caused major transcriptional changes (Supplementary Figure 4f and Supplementary Table 4). However, we successfully observed previously detected pathways, as well as potential side effects of this drug in ctrl and *MSL3* patient cells[38] (Figure 4d and Supplementary Table 4). This supports that the primary dermal fibroblast cell lines used in our analysis represent a relevant model for studying the *MSL3* syndrome.

Interestingly, a large set of the DE-genes reversed their expression towards normal levels upon HDACi albeit the extent varied between different genes and patients (Supplementary Figure 4g and Supplementary Table 3). We focused our attention on the 116 genes which were fully rescued to control levels and hence did not score as DE after panobinostat treatment (Figure 4e). Among them, clinically relevant genes based on the patient phenotypes were scrutinized. One would expect a positive clinical impact by restoring a basal expression of e.g. *SOX11* (OMIM:615866), *HOXD10* (OMIM:142984), *LEP* (OMIM:164160) or *ACTG2* (OMIM:102545) based on their role in human disease and patient symptoms including cognitive performances, gait disturbance, overweight and feeding difficulties/constipation, respectively. Accordingly, KEGG pathway analysis revealed that HDACi may generally rewire

deregulated processes relevant to the disease state of *MSL3* patients (Figure 4f and Supplementary Table 3).

Hence, we wondered whether we would be able to rescue the previously observed cell cycle phenotype and migratory signature using HDACi. Indeed, genes involved in G0 arrest, e.g. *GAS1*, were rescued towards the expression levels in ctrl cells upon HDACi (Supplementary Figure 4h), while displaying a slight increase in G1/S arrest. Moreover, in agreement with the activation of the respective GO terms (Figure 3h), we had observed that *MSL3* patient-derived fibroblasts migrate faster than ctrl cells. We next assessed whether the positive influence of HDACi on gene expression would also lead to a beneficial outcome on a cellular level. Strikingly, upon HDACi treatment we could restore the phenotype of *MSL3* mutant cells, as they now adopted the migratory behavior of the control cells (Figure 4g, h and Supplementary Figure 4i). Therefore, HDACi have the potential to both molecularly as well phenotypically alleviate at least some of the defects in *MSL3* mutant cells, thus providing a promising therapeutic avenue for MSL3 patients.

Discussion

In summary, we report a novel syndrome, where loss of *MSL3* is associated with a pronounced loss of bulk H4K16ac. The patients presented in our study display a complex phenotype including neurodevelopmental delay, with a recurrent facial dysmorphism and a progressive neurological disorder. This indicates that MSL-mediated H4K16ac in humans is involved in the regulation of complex developmental processes, for example those occurring in the brain[5,39,40]. This is intriguing as it is in contrast to the sex-specific function of MSL3 during dosage compensation in flies, suggesting a more context specific role in mammals. Given the conserved role of H4K16ac in chromatin decompaction from yeast to humans, it remains to be demonstrated why its physiological roles evolved to be so organism specific. Our study also raises a general question on the role of H4K16ac, as *MSL3* patients survive to adulthood compared to the early embryonic lethality of the *Mof* mutant in mouse[41–44]. An important point to consider in this context is that although MSL3 mutations lead to a severe loss of H4K16 acetylation, remaining H4K16ac may be sufficient to allow developmental progress to a certain extent, since *MOF* is still expressed in these cells. While no *MOF* variants have been reported in any

developmental disorder yet, *de novo* heterozygous loss-of-function variants of *KANSL1* were reported to cause Koolen-de Vries syndrome (KdVS)[45–47]. *KANSL1* encodes a scaffold component of the MOF-KANSL complex. However, the disease characteristics of KdVS are symptomatically and molecularly distinct from *MSL3* patients with regards to features of facial dysmorphism, specifically palpebral fissures orientation, nose tip and mouth shape, visceral and brain abnormalities, which are absent in the reported *MSL3* patients. Although KdVS patients also display neuromuscular disorders, the progressive gait disturbance associated with *MSL3* syndrome may be a distinctive feature. Given the multifaceted and widespread roles of MOF in processes such as metabolism[44], autophagy[48] and stress response[43], we believe that the *MSL3* mutations presented here may serve as an elegant separation-of-function allele, allowing us to dissect the role of altered H4K16ac deposited by the MOF-MSL versus the MOF-KANSL complex during human development.

The *MSL3* syndrome affects both hemizygous males and heterozygous females with comparable symptoms, which is difficult to reconcile with classical models of X chromosomal inheritance[49]. More recently, it has been appreciated that X-linked disorders can display intermediate penetrance, for example due to varying degrees of XCI in females[50]. This includes variations in escape genes, such as *IQSEC2*[51], or examples like *XLH* or *E1α*, where similar phenotypes can be observed in both heterozygous females and hemizygous males. For *MSL3* this can be presumably linked to its property of being a facultative escapee, where various degrees of XCI and escape might provide an explanation for the phenotypic and molecular differences amongst female *MSL3* patients. This is notable for the mother of p16, which is a carrier of the disease-causing allele, but is only mildly affected. The nature of the *MSL3* variants, as well as environmental effects modulating epigenome plasticity might additionally contribute to the phenotypic outcome of the syndrome. A comprehensive assessment of the allele-specific expression of *MSL3* in different types of tissues and the establishment of alternative cellular and animal models will be required to dissect the apparently very complex regulation and phenotypic outcome in *MSL3* syndrome in both sexes.

Lastly, the positive outcome of HDACi partially alleviating migratory and cell cycle defects provides a preliminary, yet promising ground for future studies, especially in light of the degenerative nature of the *MSL3* syndrome. More efforts will need to be

undertaken to potentially ameliorate the symptomatology and delay the likely myelopathy-caused hyperreflexia, for which HDACi have already shown positive results[52]. The possibility of alleviating at least some of the symptoms of *MSL3* patients through anti-cancer drug repurposing, provides yet another case where epigenetic therapy using HDACi may pave ways for treatments of other diseases affecting epigenetic modifiers. Taken together, our characterization of this novel syndrome not only provides insights into the developmental roles of MSL-associated H4K16ac in humans but also underscores the general relevance of studying sex-specific aspects related to X-linked diseases.

Accession Codes / Data availability

RNA-seq data have been deposited to the Gene Expression Omnibus under the accession number GSE102250.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[21] with the dataset identifier PXD009317.

The damaging variants reported as disease-causing in this article were deposited in ClinVar database under study accession SUB2871008.

Acknowledgments

We are grateful to the patients and their families for their cooperation and interest in the study. We thank N. Iovino, B. Sheikh and I. Ilik for critical reading of the manuscript. We also thank C. Pessoa Rodrigues and A. Karoutas for technical help, insightful discussion and advices. We thank V. Bhardwaj for advice and consulting on RNA-seq analysis and S. Kübart, A. Schröer, J. Wirth and H.-G. Nothwang for their help with inversion breakpoint mapping. We thank L. Wells for patient recruitment and clinical data collection. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the

Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. This study makes use of DECIPHER (http://decipher.sanger.ac.uk), which is funded by Wellcome. Sequencing for Patient 12 was provided by the Center for Mendelian Genomics at the Broad Institute of MIT and Harvard and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung and Blood Institute grant UM1 HG008900 to Daniel MacArthur and Heidi Rehm. This work was supported by CRC992, CRC1140 and CRC746 awarded to A.A., the work was supported by the council of Burgundy, the German Human Genome Program (DHGP, grant number 01KW99087), and the National Genome Research Network (NGFN, project numbers 01GR0105 and 01GS08160) awarded to V.M.K. and A.R., respectively. C.I.K.V. was supported by a Human Frontier Science Program (HFSP) long-term fellowship (000233/2014-L).

Contributions

The project was conceived and the experiments were planned by M.F.B., A.A., J.T.; the review of phenotypes and the sample collection was performed by J.M., I.B., M.K., H.G.B., O.V., K.Õ., B.M., T.Y.T., S.V., K.L., A.R., D.S.J., A.F., V.M.C., R.B.F., V.M.K., A.B., L.F., S.E.C.D., J.T., C.T.R., I.vdB., M.T.C., M.W.; fibroblast isolation, tissue section preparation and processing, cell culture, protein analysis, FACS, immunostaining, microscopy, drug treatment experiments, ChIP, RNA expression analysis and IPA pathway analyses were performed by M.F.B.; A.B. initiated data-sharing and compiled patient's clinical data; G.S. performed deep-sequencing data analysis and manuscript editing; C.I.K.V. contributed to design, analysis and interpretation of experiments and data; T.A. performed cloning of MSL3 expression constructs and co-immunoprecipitations; W.G.S. and G.M. performed LC-MS sample preparation and analysis; T.R. purified recombinant proteins, performed HAT assays and in silico modelling. Ethical consultation was provided by L.F.; data analysis for

exome and frequency calculation was performed Y.D., S.P., J.T., A.B., C.G., R.P.; the manuscript was written and edited by C.I.K.V., M.F.B., A.A., J.T.; all authors reviewed, edited and approved the paper.

Competing Financial Interests

The authors declare no competing financial interests

References

1. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. *Nat. Rev. Genet.* **14,** 347–359 (2013).
2. Ropers, H.-H. & Hamel, B. C. J. X-linked mental retardation. *Nat. Rev. Genet.* **6,** 46–57 (2005).
3. Lubs, H. A., Stevenson, R. E. & Schwartz, C. E. Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* **90,** 579–590 (2012).
4. Bjornsson, H. T. The Mendelian disorders of the epigenetic machinery. *Genome Res.* **25,** 1473–1481 (2015).
5. Tapias, A. & Wang, Z.-Q. Lysine Acetylation and Deacetylation in Brain Development and Neuropathies. *Genomics Proteomics Bioinformatics* **15,** 19–36 (2017).
6. Lee, J.-H., Hart, S. R. L. & Skalnik, D. G. Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* **38,** 32–38 (2004).
7. Hsieh, J., Nakashima, K., Kuwabara, T., Mejia, E. & Gage, F. H. Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 16659–16664 (2004).

8. Mews, P. *et al.* Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. *Nature* **546,** 381–386 (2017).
9. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15,** 703–708 (2014).
10. Shogren-Knaak, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311,** 844–847 (2006).
11. Keller, C. I. & Akhtar, A. The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond. *Curr. Opin. Genet. Dev.* **31,** 1–11 (2015).
12. Philippakis, A. A. *et al.* The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum. Mutat.* **36,** 915–921 (2015).
13. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536,** 285–291 (2016).
14. Smith, E. R. *et al.* A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol. Cell. Biol.* **25,** 9175–9188 (2005).
15. Taipale, M. *et al.* hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol. Cell. Biol.* **25,** 6798–6810 (2005).
16. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348,** 648–660 (2015).
17. Kadlec, J. *et al.* Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.* **18,** 142–149 (2011).
18. Zhao, X. *et al.* Crosstalk between NSL histone acetyltransferase and MLL/SET complexes: NSL complex functions in promoting histone H3K4 di-methylation activity by MLL/SET complexes. *PLoS Genet.* **9,** e1003940 (2013).

19. Gelbart, M. E., Larschan, E., Peng, S., Park, P. J. & Kuroda, M. I. Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat. Struct. Mol. Biol.* **16,** 825–832 (2009).
20. Akhtar, A. & Becker, P. B. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. *Mol. Cell* **5,** 367–375 (2000).
21. Chelmicki, T. *et al.* MOF-associated complexes ensure stem cell identity and Xist repression. *Elife* **3,** e02024 (2014).
22. Tukiainen, T. *et al.* Landscape of X chromosome inactivation across human tissues. *Nature* **550,** 244–248 (2017).
23. Cai, Y. *et al.* Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. *J. Biol. Chem.* **285,** 4268–4272 (2010).
24. Morales, V., Regnard, C., Izzo, A., Vetter, I. & Becker, P. B. The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex. *Mol. Cell. Biol.* **25,** 5947–5954 (2005).
25. Kim, D. *et al.* Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat. Struct. Mol. Biol.* **17,** 1027–1029 (2010).
26. Vaquero, A. *et al.* SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis. *Genes Dev.* **20,** 1256–1261 (2006).
27. Bonenfant, D. *et al.* Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. *Mol. Cell. Proteomics* **6,** 1917–1932 (2007).
28. McManus, K. J. & Hendzel, M. J. The relationship between histone H3 phosphorylation and acetylation throughout the mammalian cell cycle This paper

is one of a selection of papers published in this Special Issue, entitled 27th International West Coast Chromatin and Chromosome Conference, and has undergone the Journal's usual peer review process. *Biochem. Cell Biol.* **84,** 640–657 (2006).

29. Conrad, T. & Akhtar, A. Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. *Nat. Rev. Genet.* **13,** 123–134 (2012).
30. Conrad, T. *et al.* The MOF chromobarrel domain controls genome-wide H4K16 acetylation and spreading of the MSL complex. *Dev. Cell* **22,** 610–624 (2012).
31. Kasowski, M. *et al.* Extensive variation in chromatin states across humans. *Science* **342,** 750–752 (2013).
32. De Ruijter, A. J. M., Van Gennip, A. H., Caron, H. N., Stephan, K. & Van Kuilenburg, A. B. P. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem. J* **370,** 737–749 (2003).
33. Marks, P. A. & Breslow, R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* **25,** 84–90 (2007).
34. Bradner, J. E. *et al.* Chemical phylogenetics of histone deacetylases. *Nat. Chem. Biol.* **6,** 238–243 (2010).
35. Coni, S. *et al.* Selective targeting of HDAC1/2 elicits anticancer effects through Gli1 acetylation in preclinical models of SHH Medulloblastoma. *Sci. Rep.* **7,** 44079 (2017).
36. Lillico, R., Sobral, M. G., Stesco, N. & Lakowski, T. M. HDAC inhibitors induce global changes in histone lysine and arginine methylation and alter expression of lysine demethylases. *J. Proteomics* **133,** 125–133 (2016).

37. Schölz, C. *et al.* Acetylation site specificities of lysine deacetylase inhibitors in human cells. *Nat. Biotechnol.* **33,** 415–423 (2015).
38. Tzogani, K. *et al.* The European Medicines Agency Review of Panobinostat (Farydak) for the Treatment of Adult Patients with Relapsed and/or Refractory Multiple Myeloma. *Oncologist* (2017). doi:10.1634/theoncologist.2017-0301
39. Volmar, C.-H. & Wahlestedt, C. Histone deacetylases (HDACs) and brain function. *Neuroepigenetics* **1,** 20–27 (2015).
40. Dulac, C. Brain function and chromatin plasticity. *Nature* **465,** 728 (2010).
41. Thomas, T., Dixon, M. P., Kueh, A. J. & Voss, A. K. Mof (MYST1 or KAT8) is essential for progression of embryonic development past the blastocyst stage and required for normal chromatin architecture. *Mol. Cell. Biol.* **28,** 5093–5105 (2008).
42. Gupta, A. *et al.* The mammalian ortholog of Drosophila MOF that acetylates histone H4 lysine 16 is essential for embryogenesis and oncogenesis. *Mol. Cell. Biol.* **28,** 397–409 (2008).
43. Sheikh, B. N. *et al.* MOF maintains transcriptional programs regulating cellular stress response. *Oncogene* **35,** 2698–2710 (2016).
44. Chatterjee, A. *et al.* MOF Acetyl Transferase Regulates Transcription and Respiration in Mitochondria. *Cell* **167,** 722–738.e23 (2016).
45. Koolen, D. A. *et al.* A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat. Genet.* **38,** 999–1001 (2006).
46. Sharp, A. J. *et al.* Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat. Genet.* **38,** 1038–1042 (2006).

47. Shaw-Smith, C. *et al.* Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat. Genet.* **38,** 1032–1037 (2006).
48. Füllgrabe, J. *et al.* The histone H4 lysine 16 acetyltransferase hMOF regulates the outcome of autophagy. *Nature* **500,** 468–471 (2013).
49. Morgan, T. H. SEX LIMITED INHERITANCE IN DROSOPHILA. *Science* **32,** 120–122 (1910).
50. Dobyns, W. B. *et al.* Inheritance of most X-linked traits is not dominant or recessive, just X-linked. *Am. J. Med. Genet. A* **129A,** 136–143 (2004).
51. Zerem, A. *et al.* The molecular and phenotypic spectrum of IQSEC2-related epilepsy. *Epilepsia* **57,** 1858–1869 (2016).
52. Lezin, A. *et al.* Histone deacetylase mediated transcriptional activation reduces proviral loads in HTLV-1 associated myelopathy/tropical spastic paraparesis patients. *Blood* **110,** 3722–3728 (2007).

Figure Legends

Figure 1

Mutations in *MSL3* cause a novel neurodevelopmental disorder

(**a**) *MSL3* patient photographs. Grey circles indicate intra-genic variants, yellow whole-gene deletions, stars donated skin biopsies. Written consent for the use of photographs was obtained from the parents of affected individuals. (**b**) Human Phenotype Ontology heatmap of *MSL3* patient common clinical features. Blank boxes represent either absent or not reported symptoms. (**c**) *MSL3* locus and identified variants highlighted in red. p1-13 variants occur within *MSL3* (top), p13/p14 deletions span *ARHGAP6*, *MSL3* and *FRMD4* (middle). p16 chromosomal inversion and breakpoints (green) are between Xp22.2 and Xq28 (bottom). (**d**) Annotation of the pathogenic variants on the primary protein structure. Grey lines represent predicted proteins, where most SNVs, except for p.(L308P) shown as green dot, result in MRG-domain truncation. (**e**) Cropped H4K16ac and H3 immunoblot of a histone acetyltransferase (HAT) assay performed on nucleosomal substrate with the indicated proteins (fl=full length (2-458), HAT domain (174-458)). The right panel (lanes 2-5) is a coomassie-stained gel of the assayed proteins. The experiment was repeated twice with similar results. **(f)** Modelled surface representation of the MSL1-MSL3 complex (PDB: 2YON, MSL1:yellow, MSL3:grey/purple). Missing parts in p1 are highlighted in purple (left) or removed (right). (**g**) Cartoon representing the MSL1 (575-585, yellow) and MSL3 (253-270/459-484, grey) interaction interface with hydrogen bonds in black. α-helices5/6 missing in p1 are shaded blue. (**h**) Flag-co-immunoprecipitations (IP) of MSL3 full-length (fl), MSL3 variants (p2, p1 and p3) GFP IP-control with cropped immunoblots. MSL3 and MSL1-HA expression vectors were transiently co-transfected into HEK293 cells. DHX9 serves as input loading control. The experiment was repeated 4 times with similar results.

Figure 2

Bulk reduction of H4K16 acetylation in *MSL3* patient cells

(**a**) Immunohistochemistry of FFPE-skin sections for H4K16ac (white, MERGE in red) and E-cadherin (green). DAPI staining in blue. Asterisks highlight cells with reduced H4K16ac staining. Scale bar=20 µm. The staining was repeated twice with similar results. (**b**) as in a) but for MOF (white) **(c)** Immunostaining of ctrl, p1, p2 and

p14-derived primary dermal fibroblasts (HDFs) for H4K16ac (red) and MSL3 (white). Scale bar=10 µm. The staining was repeated twice (*n*=2 sections per slide). **(d,e).** Quantification of staining intensities for H4K16ac (d) and H3K27me3 (e) of one representative experiment in ctrl, p1, p2 and p14 HDFs. Each data point corresponds to the intensity of a single cell with the center line representing the average±SEM. p-values were determined by Ordinary One-way ANOVA followed by Bonferroni multiple corrections. **(f)** Cropped immunoblot of HDF nuclear extracts. 100% and 50% dilutions were loaded. The experiment was reproduced three times. **(g)** Volcano plot of indicated histone modification abundance displayed as log2FC (x-axis, p1, p2 and p14 HDFs (n=6, 2 for each patient) in comparison to ctrl HDF (n=3)) versus the statistical significance (y-axis, two-sided moderated *t*-test) as determined by LC-MS. Vertical dashed lines represent 4-fold changes and a horizontal line shows a q-value $-\log_{10}1.2$, Benjamini-Hochberg-corrected (BH) FDR=0.05. Circles in red=acetylated sites, violet=monomethylated sites, light-blue=propionyl-methylated sites, dark blue=trimethylated sites. Detailed results of statistical analyses in **(d)-(e)** are provided in Supplementary Table 5, for **(g)** in Supplementary Table 2.

Figure 3

Loss of *MSL3* leads to transcriptome misregulation of developmental pathways

(a) (Left) Heatmap representing z-scores of DE-genes (p1/p2/p14 vs ctrl) obtained by RNA-seq (FDR 0.05). (Right) MACS2 ChIP-seq[21] peaks for mMSL1, mMSL2 and mMOF at mouse orthologues of *MSL3* patient DE-genes illustrated with green, white (no peak), or grey boxes (no orthologue). **(b)** DE-genes pie chart. **(c)** Box plot displaying DESeq2 log2FC SEM for the DE-downregulated (n=196) versus the DE-upregulated genes (n=323). The p-value was calculated using a two-sided Welch Two Sample *t*-test. **(d)** ChIP-seq peak frequency plot for mouse mMSL2 or mKANSL3 at unaffected/not-DE versus misregulated/*MSL3*-DE-genes in patients. P-values for ChIP-peak over-representation or under-representation within the groups were calculated using a one-sided Fisher's test. **(e)** Dot plots (center line representing the average±SEM) of expression and H4K16ac ChIP qPCR analyses. Expression levels were determined in ctrl and patients (left) and upon *MSL3* knock-down (KD, middle) in male (♂) and female (♀) ctrl HDFs. H4K16ac ChIP-qPCR enrichment values were calculated relative to input and expressed as fold-change

over *KLK3*. (Right) Genome-browser snapshots of mMSL2 ChIP-seq enrichment at corresponding mouse orthologues **(f)** RT-qPCR expression analysis displayed as dotplots with center line representing the average±SEM. Expression levels were normalized to *RPLP0* and expressed relative to ctrl. p-values were determined by Ordinary One-way ANOVA followed by Bonferroni multiple corrections. **(g)** IPA z-score plot for Disease and Function signature of KEGG pathways (padj<0.005 cutoff). BH multiple testing correction was applied (*n*=519 DE-genes). **(h)** Bubble plot representing perturbation pathway analysis for molecular and cellular functions. Bubble size=Molecule abundance, Color=IPA z-score. The most representative sub-pathway by number of molecules was selected to represent the full signature (*n*=477 transcripts). **(a)-(f)** In dotplots, each dot represents an independent experiment in HDFs. Further details and statistical test values of each plot are provided in Supplementary Table 5.

Figure 4

HDAC inhibition alleviates cellular and molecular phenotypes of *MSL3* patient cells

(**a**) *MSL3* and empty expression vector (EV) were transfected in ctrl and p1/p14 HDFs. Dot plots depict RNA levels of each independent experiment normalized to *RPLP0* relative to EV with center line representing the average (*n*=2). **(b)** Immunofluorescence for H4K16ac (red) and MSL3 (green) upon treatment with 4 µM LBH-589. Scale bar=5 µm. The staining was reproduced twice. (**c**) Quantification of (b). Data points represent H4K16ac intensities of individual cells from one representative experiment with center line representing the average±SEM. p-values were determined by Ordinary One-way ANOVA followed by Bonferroni multiple corrections. Detailed results of statistical analyses are shown in Supplementary Table 5. **(d)** Heatmap of activation z-scores (z) obtained from IPA (Comparison Analysis, Disease and Function signature, (*n*=3391 ctrl, *n*=2740 p1, *n*=2487 p2, *n*=2737 p14 with padj<1E-4 DE-genes upon LBH-589 treatment (2 µM)). Activated pathways were plotted for |±z|>2.5 and p-value cutoff of $\log_{10}1.5$. Bold highlights pathways impacting on patient cell physiology. **(e)** Heatmap representing z-scores of DE-genes reverting to not-DE anymore upon LBH-589 treatment (*n*=109). Five disease-relevant transcripts are highlighted. **(f)** IPA comparison analysis of KEGG pathway z-scores in genes rescued upon LBH-589 (*n*=109 with expression cutoff of

log ratio>|±0.2|). **(g)** Representative DIC images of HDFs at 0 and 24h after creating a “scratch” (gap area). **(h)** Quantification of the % gap area at 0h, 6h, 24, 48 hours for ctrl and patient HDFs grown with or without LBH-589 (2 μM). The center line represents the average±SEM (*n*=3).

Table 1

	Genomic position hg19	**cDNA** NM_078629.3	**Protein** Q8N5Y2-1	**Inheritance**
Patient 1	chrX:g.11790375G>T	c.1381+1G>T	NA	*de novo*
Patient 2	chrX:g.11790367_11790374del	c.1374_1381del	p.(Leu459Glufs*13)	*de novo*
Patient 3	chrX:g.11783713C>T	c.1036C>T	p.(Gln346*)	*de novo*
Patient 4	chrX:g.11783600T>C	c.923T>C	p.(Leu308Pro)	*de novo*
Patient 5	chrX:g.11790365C>T	c.1372C>T	p.(Arg458*)	*de novo – Parental mosaicism*
Patient 6	chrX:g.11790365C>T	c.1372C>T	p.(Arg458*)	*de novo – Parental mosaicism*
Patient 7	chrX:g.11780349_11780350delAC	c.566_567delAC	p.(Tyr189Leufs*3)	*de novo*
Patient 8	chrX:g.11783695delG	c.1018delG	p.(Ala340Leufs*9)	*de novo*
Patient 9	chrX:g.11783615dupC	c.938dupC	p.(Leu314Phefs*18)	*de novo*
Patient 10	chrX:g.11783802_11783818dup17	c.1125_1141dup17	p.(Met381Argfs*30)	*de novo*
Patient 11	chrX:g.11783850T>G	c.1171+2T>G	N/A	*de novo*
Patient 12	Del exon 10	.	.	*de novo*
Patient 13	chrX:g.11781990C>T	c.841C>T	p.(Gln281*)	*de novo*
Patient 14	chrX:g.11633731_11797224del	.	.	*de novo*
Patient 15	chrX:g.11600773_12249902del	.	.	NA – Father not tested, absent in the Mother
Patient 16	Inv(X)(p22.2q28)	.	.	Inherited from the mother – *de novo* in the mother

MSL3 variants identified in the 16 reported individuals of this study.

NA: Not available.

Online Methods

Subjects, whole-exome sequencing and variant validation

Sixteen individuals from 15 unrelated families with a disease-causing variant in *MSL3* were included in this study after obtaining informed consent. The local ethic committee of the Dijon University Hospital approved this study (2016-A01347-44). The patients were followed-up by clinical geneticists from France, United Kingdom, the Netherlands, Belgium, Denmark, Germany, the US, Estonia and Australia. Thirteen variants were identified by whole-exome sequencing for the diagnostic of an undiagnosed rare disorder associated with multiple congenital anomalies and syndromic intellectual disability (ID), two deletions were identified by array-CGH and SNP-array, and the X chromosome inversion breakpoint in *MSL3* was mapped by FISH and cloning of the junction fragment. The clinical observations were gathered through the Matchmaker Exchange initiative.

Blood or DNA samples from affected individuals and their parents, and informed consent were obtained from all participants in accordance with site-specific institutional boards of each country. Dedicated consents for the collection of a skin biopsy was obtained.

De novo variants were analyzed by PCR and direct Sanger sequencing using DNA from patients and their parents. Primer sequences and detailed methods are available on request. In the family with two affected siblings (p5 and p6) and suspected parental somatic mosaicism, confirmation and estimation of the percentage of cells bearing the variant was performed using DNA extracted from peripheral blood and saliva of all members of the family (the two patients and parents). Targeted sequencing for the variant position was performed on a MiSeq instrument (Illumina), at a depth of 1706 and 1398 folds for the variant position in the father and mother, respectively. Genotyping identified 12 and 6 variant alleles, resulting in an allelic balance of 0.007 and 0.004, respectively. In p1, the splice variant analysis was conducted on total RNA extracted from peripheral blood with TRIzol (Life Technologies). Sanger sequencing confirmed the splicing defect. Written consent was obtained from the parents and/or guardians of the affected individuals both for participation in the study and for use of the clinical photographs in this report.

Enrichment of de novo *MSL3* variants in the matchmaker cohort

The statistical significance after Benjamini–Hochberg correction for enrichment of functional and/or LoF *de novo* mutations in the matchmaker cohort was estimated as previously described in[53]. The gene-specific mutation rates were published in[54]. Because *MSL3* is X-linked, we assumed that the matchmaker cohort had an equivalent proportion of males and females (5000 each, 15000 alleles). The uncorrected p-value was corrected for multiple testing taking 18280 genes to apply the Benjamini-Hochberg False Discovery Rate correction.

Exome analysis

Each collaborating team followed a published methodology for sample and library preparation, sequencing data production, analysis and interpretation[55–59].

Parental segregation analysis

Genomic DNA was amplified by Polymerase Chain Reaction (PCR) using HotStarTaq PCR kit (Qiagen). PCR products were purified by Agencourt CleanSEQ system (Beckman Coulter) and sequenced with the BigDye Terminator Cycle Sequencing kit, v3.1 (Applied Biosystems) on ABI 3730 sequencer (Applied Biosystems). Sequence data were analyzed using Mutation Surveyor v4.0.9 (Softgenomics).

MiSeq cDNA analysis

Total RNA was extracted from peripheral blood with TRIzol (Life Technologies). The integrity of the RNA was assessed on 1.2% agarose gel, and the concentration and purity was determined by optical densitometry. cDNA was generated with the QuantiTect Reverse Transcription Kit (Qiagen) and PCR amplified (Supplementary Table 5) with PrimeStar GXL (Takara). Nextera XT DNA Preparation kit (Illumina) was used to create a cDNA library and sequenced on a Miseq instrument (Illumina). Resulting data were aligned to the human reference genome (GRCh37/hg19) using STAR2 (v2.5.3)[60]. Aligned reads were sorted using Picard (v2.17). Splice junctions and transcript organization were visualized using the sashimi plot of the Integrative Genomics Viewer.

Haplotype

The *MSL3* locus was amplified from genomic DNA by long range PCR using PrimeStar GXL (Takara). PCR products were pooled for each case. Nextera XT DNA Preparation kit (Illumina) was used to create a tagged and fragmented DNA library, which was sequenced on a MiSeq (Illumina). Resulting data were aligned to the reference genome (GRCh37/hg19) using BWA (Burrows-Wheeler Aligner; v0.7.6). The Genome Analysis Toolkit (GATK; v2.1-10) enabled indel realignment and base quality score recalibration. Variants with a quality score >30 and alignment quality score >20 were annotated with SeattleSeq SNP Annotation (v137). Haplotype was determined using HaplotypeCaller from VCF files.

Protein modeling

The complex structure of the MSL1-MSL3 complex (PDB: 2Y0N) was analyzed and visualized with Pymol (v1.7.4)

Cloning for protein expression and purification

The human $MOF_{174-458}$ and MOF_{2-458} (Uniprot: Q9H7Z6-1) coding sequences (CDS) were cloned in a modified pET15b vector (His_{10}tag). $MSL1_{470-520}$ or $MSL1_{471-616}$ (Uniprot: Q68DK7-1) and $MOF_{174-458}$ (MCS2) CDS were cloned in a pACYC coexpression vector (Novagen, MCS1: $MSL1_{470-520}$ or $MSL1_{471-616}$; MCS2: $MOF_{174-458}$). The CDS of $MSL3_{167-517}$ (Uniprot: Q8N5Y2-1) was cloned in MCS1 of a modified pRSFDuet vector. Proteins were expressed in LB using Rosetta2 pLysS ($MOF_{174-458}$, MOF_{2-458}) or BL21(DE3) cells ($MOF_{174-458}$-$MSL1_{470-520}$, $MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$). Expression was induced with IPTG (0.1-1 mM, OD_{600} 0.5-0.7, 18 °C overnight). After lysis by sonication ($MOF_{174-458}$, $MOF_{174-458}$-$MSL1_{470-520}$, $MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$: 50 mM HEPES, 500 mM NaCl, 20 mM imidazole, 5 mM beta-mercaptoethanol, 0.01 % (w/v) taurodeoxycholate, 1 EDTA-free cOmplete (Roche), pH 7.5; MOF_{2-458}: 50 mM HEPES, 100 mM NaCl, 20 mM imidazole, 5 mM beta-mercaptoethanol, 0.01 % (w/v) taurodeoxycholate, 1 cOmplete (Roche), pH 7.5), and clarification by centrifugation, the supernatant was applied to a HisTrap HP column (GE Healthcare). This was washed with lysis buffer (for $MOF_{174-458}$, $MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$ containing 2000 mM NaCl) before elution with an imidazole gradient (20-300 mM). Protein containing fractions were pooled.

After buffer exchange (50 mM HEPES, 100 mM NaCl, 20 mM imidazole, 5 mM beta-mercaptoethanol, 0.01 % (w/v) taurodeoxycholate, pH 7.5), the solutions containing MOF_{2-458} and the trimeric MOF complex were applied to a Heparin HP column (GE Healthcare) and eluted with a NaCl gradient (100-2000 mM NaCl). Protein-containing fractions were pooled and concentrated. Finally, $MOF_{174-458}$, MOF_{2-458}, $MOF_{174-458}$-$MSL1_{470-520}$ or $MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$ were applied to a Superdex S75 ($MOF_{174-458}$, MOF_{2-458}, $MOF_{174-458}$-$MSL1_{470-520}$) or S200 ($MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$) gel filtration column (25 mM HEPES, 500 mM NaCl, pH 7.5). Protein-containing fractions were pooled, concentrated, flash-frozen and stored at -80 °C. Chromatography was performed on Äkta purifier systems. The purity of all assayed proteins and complexes was > 90 %. Protein concentration was determined by Bradford assay.

Nucleosome reconstitution, histone acetyltransferase assays

Nucleosomal core particles (NCP) were reconstituted using a PCR-amplified 147 bp DNA fragment[61] with 5' Cy5 and 3' Biotin and *Xenopus laevis* histone octamers applying salt dialysis[62]. NCP quality was assessed using agarose gel electrophoresis. For histone acetyltransferase assays, 100 nM $MOF_{174-458}$, MOF_{2-458}, $MOF_{174-458}$-$MSL1_{470-520}$, or $MOF_{174-458}$-$MSL1_{471-616}$-$MSL3_{167-517}$ were mixed with AcCoA (Sigma, final concentration 33 µM), and nucleosomes (0.54 µg) in assay buffer (10 mM Tris, 200 mM NaCl, 1 mM EDTA, pH 8.0) and incubated for 1 h (37 °C, 600 rpm). The reaction was stopped by addition of 4xSDS sample buffer followed by SDS-PAGE and immunoblot.

Cell culture, drug treatment and transfections

Primary HDFs were isolated from 3 mm skin biopsies using 0.25% Collagenase type I, 0.05% DNase I, 20% FCS high glucose DMEM (Gibco) overnight at 37°C. Dermal fibroblasts were let to attach and cultured in DMEM supplemented with 10% FCS pen-strep. All experiments were performed at early passages (2 to 8). LBH-589 (Selleck Chemicals Co. Ltd.), MGCD0103 (Toronto Research Chemicals), SAHA (Adipogen), TSA were dissolved in DMSO. For most experiments, prior to experimental treatment, cells were starved overnight in HBSS medium (24020133, Thermo Fisher). Flp-In 293 T-REx cells (R78007, Thermo Fisher) were maintained in DMEM-Glutamax supplemented with sodium pyruvate, glucose and 10% FBS. They

were also maintained in zeocin- and blasticidin-containing medium according to the manufacturer's protocol. All transgenes used in immunoprecipitation experiments were cloned into pCDNA5-FRT/To (V6520-20, Thermo Fisher) with a C-terminal 3XFlag–HBH. *MSL1* CDS was cloned with a C-terminal 3xHA-Avi tag. This vector was co-transfected with *MSL3* WT (full length=fl), truncating mutants or with GFP CDS (as a IP negative control) in 1:1 ratio (500 ng each on 6 well plates for nuclear FLAG-IPs) with a 3:1 Lipid (µl of Lipofectamine 2000) to DNA (µg) ratio. 6 hours after transfection, transgenes were induced with 0.1 µg/ml Doxycyline for 16 hours before lysate preparation. The hsMSL3-3xFlag-HBH CDS was subcloned into a custom pCAG-MCS-BGH(polyA), pSV40-BlasticidinR-SV40(polyA), pSV40-mCherry-SV40(polyA) vector created in multiple steps from SV40 promoter containing backbone pMB1610 (addgene 65853). The vector and map can be provided upon request. For HDF transfections, 0.1 µg plasmid was used with Lipofectamine 3000 on $1x10^5$ cells in OPTI-MEM followed by selection with 30 µg/ml blasticidin-HCl (Invitrogen A111139-03) overnight. Cells were then grown in fresh media for 24h before RNA extraction.

FLAG-immunoprecipitation (IP)

Nuclei were isolated with 0.5% NP-40 in isotonic lysis buffer (10 mM Tris HCl, pH 7.5, with 2 mM $MgCl_2$, 3 mM $CaCl_2$, 0.3 M Sucrose). Nuclear pellet was extracted using HMGT buffer (25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 10% glycerol, 0.2% Tween-20) containing 420 mM KCl. The lysate was clarified by centrifugation before slowly reducing KCl to 150 mM by dilution. FLAG-M2 beads (Sigma 1804) were used for overnight IP followed by HMGT buffer (150 mM KCl) washes. Proteins were eluted by incubating 10 minutes at 70ºC in 2xSDS sample buffer.

***MSL3* knock-down in HDFs**

Silencer Select (Ambion-Thermo Fisher) siRNA ID s21530 and control siRNA were used at a final concentration of 50 nM for 32 hours before RNA extraction.

Expression analysis in HDFs and RT–qPCR

Total RNA was prepared from 12-well plate cultured cells using TRIzol (Thermo Fisher) and purified with Direct-zol™ RNA MiniPrep and MicroPrep kits (Zymo Research). cDNA samples were synthesized using Maxima Reverse Transcriptase

(K1671, Thermo Fisher) from 20-50 ng of total RNA. RT–qPCR reactions were performed on a LightCycler® 480 Multiwell Plate 384 (Roche) using SYBR Green Master Mix (Roche). Primer efficiency was tested using serial dilutions.

Immunofluorescence and Immunohistochemistry

HDFs were seeded as required for the experimental setup and fixed in 4% formaldehyde in µ-Slide 8-well on ice. FFPE-section and HDFs immunostainings were performed according to published procedures[63] using MSL3 antibody (USBI129933, MaxPab, 1:200) as primary together with H4K16ac (07-329, Millipore, 1:500), H3K27me3 ((C36B11) 9733, Cell signalling, 1:250), E-cadherin (610181, BD Bioscience, 1:200), KAT8 [EPR15803] (ab200660, Abcam, 1:500) antibodies.

Multi-stacked and tailed images were acquired on a Spinning disk confocal microscope (Zeiss). Fluorescence intensity was measured on orthogonal projections from the same stack number using ImageJ (1.47v) and corrected for background. Fluorescently-labeled secondary antibodies (Invitrogen) were used at 1:300 dilution: goat-anti-mouse-Alexa594 (A-11032), goat-anti-rabbit-Alexa594 (A-11037), goat-anti-rabbit-Alexa488 (A-11034), goat-anti-mouse-Alexa488 (A-11001).

Immunoblotting

Cells were fractionated using a subcellular fractionation kit (78840, Thermo Fisher), the buffers supplemented with cOmplete (Roche) and sodium butyrate (19-137, MerckMillipore). Protein concentration was determined using Qubit protein assay reagent. Samples were denatured at 95 °C for 10 minutes in Pierce™ Lane Marker Reducing Sample Buffer, resolved by NuPAGE Bis-Tris denaturing conditions, transferred to 0.22 µm PVDF membranes followed by blocking in 5% nonfat milk in PBS (0.3% Tween) before incubation with antibodies: MSL3 (USBI129933, MaxPab, USBiologicals, 1:1000 in BSA), H3 (61475, Active Motif, 1:5000), H4K16ac (39167, Active Motif, 1:2000; 07-329, Millipore, 1:2000), Flag (M2, Sigma, 1:1000), HA (MMS-101P, Covance, 1:1000), MOF (A300-992A, Bethyl, 1:1000), DHX9 (NDH-II(E10) sc-137183, 1:2000-1:5000), H3K36me3 (ab9050, Abcam, 1:1000), acH3 (rabbit, Millipore, 1:1000), H4K20me1 (ab9051, Abcam, 1:1000), H4K8ac (61103, Active Motif, 1:1000), H3K14ac (39599, Active Motif, 1:1000), acH4 (06-598, Millipore, 1:1000), H3K27me3 (9733, Cell Signaling, 1:1000).

Mass Spectrometry sample preparation

After histone extraction (78840, Thermo Fisher), proteins were separated on a 16 % Tris-Glycine gel (Novex™) and stained with colloidal coomassie. Gel bands containing histones were cut out. After destaining, reduction and in gel alkylation, histones were propionylated in a 79:1 mixture of isopropanol and propionic anhydride (240311, Sigma) for 2h. Gel pieces were washed with MiliQ water, dried in a SpeedVac centrifuge and rehydrated in hydroxylamine followed by incubation with ammonium hydroxide solution to reestablish neutral pH. After further washes with ammonium bicarbonate-based buffer and acetonitrile, gel pieces were again dehydrated. Subsequent tryptic digestion was performed as described in[64]. Recovered peptides were desalted on STAGE tips according to[65].

LC-MS bottom-up analysis of histone peptides

Digested and desalted peptides were analyzed by LC-MS/MS using nanoflow HPLC (Easy nLC1200, Thermo Fisher) coupled with QExactive hybrid mass spectrometer (Thermo Fisher) through a nanoelectrospray ion source (Proxeon). Peptides were loaded on a 20 cm long, 75 µm wide analytical column filled in-house with 1.9 µm C18 beads (Reprosil C18, Dr. Maisch GmbH, Germany) at a flow rate of 500 nl/min. Analytical column was then equilibrated for 5 minutes (5% buffer A) which was followed by a 60 minutes long peptide elution at flow rate of 300 nl/min with a linearly increasing concentration of organic solvent (80% acetonitrile, 0.1% formic acid). QExactive was operated in positive mode, employing data-dependent acquisition of fragmentation spectra of multiple-charged peptides for their identification.

MaxQuant software (v1.6.0.16) was used for raw file peak extraction and protein identification. Peptide and protein False Discovery Rate (FDR) was set to 0.01 using a decoy database. Additionally, “match between runs” was used with matching time window of 1 minute and alignment time window of 20 minutes.

LC-MS data analysis

List of ratios of modified lysine or arginine sites derived from MaxQuant (“…sites.txt”) was imported to Perseus[66]. Data matrix was modified by applying “Expand site table” step, removing empty rows, as well as contaminants and entries from decoy database. All site intensities were log2 transformed and normalized by subtracting the mean of the whole sample from the particular histone site intensities[67]. A

moderated *t*-test was used to test for significant differences between control and patient samples. Differences with a qvalue<0.05 (Benjamini-Hochberg corrected) were considered as significant.

Paraffin embedding and sectioning

2mm human skin specimens were fixed overnight in PBS containing 4% formaldehyde. After dehydration in ethanol/PBS and Histolemon series, they were incubated in paraffin at 58°C overnight and casted into molds covered with paraffin. Blocks were sectioned using an RM2155 microtome (Leica) at 5-10 µm. Additional FFPE skin sections age-matching to ctrl were purchased from BioCat.

Hematoxylin and Eosin (H&E) staining

Paraffin sections were dewaxed and rehydrated step-wise. The samples were stained in hematoxylin solution (0.1% hematoxylin, 5% $KAl(SO_4)_2$, 0.02% KIO_3). Counterstaining was performed by incubating the slides in eosin solution (1% eosin). Slides were mounted in Permount Mounting Medium (Thermo Fisher) and analyzed using brightfield microscopy.

Cell cycle analysis by FACS

HDFs were synchronized by mitotic shake followed by serum starvation overnight. Cells were induced to cycle for 6 hours by addition of complete media. $1x10^6$ HDFs were fixed in suspension in 70% cold ethanol while gently shaking for 20 minutes at 4°C. After washing in FACS buffer (0.2% BSA in PBS), H4K16ac staining was performed at RT for 90 minutes (39167, Active Motif, 1:500; 07-329, Millipore, 1:500 in FACS buffer). This was followed by 30 minute incubation with Ki67-FITC as indicated (BD Biosciences, 556026) and 2ndary anti-rabbit (APC labeled). Propidium iodide (PI) (1 µg/ml) was used to stain nuclei. After additional washes, cells were analysed in FACS buffer using LSRII (BD, Heidelberg, Germany).

Alive singlets, based on FSC/SSC, were gated for further analysis. Three channel analyses were performed using antibodies/dyes with no fluorescence absorbance overlap. Machine Laser intensities were set up using beads. For downstream analysis in FlowJo (v10.4.1), alive cells were gated, followed by re-gating of only single cells. For cell cycle phases analyses, “Cell cycle” automatic function was run using PI.

Colorimetric Detection of Senescence-Associated β-Galactosidase

Performed as described in[68]. Pictures were taken with a conventional CCD camera on a regular inverted cell culture microscope with a 10x objective.

RNA-seq

Approximately 50 ng RNA was isolated from synchronized ctrl, p1, p2 and p14 HDFs at passage 4 and passage 7 and subjected to poly(A) RNA-seq (TruSeq® Stranded mRNA Sample Preparation kit, Illumina). Reads were mapped to the GRCh37/hg19 reference genome with HISAT2 (v2.2.0-beta)[69] using default parameters. Primary alignments were assigned to features with featureCounts (v1.4.6-p2)[70] using the GRCh37 Gencode annotation (release 19). Differential expression (DE) analysis was performed using DESeq2 (v1.18.1)[71] with default settings. We identified DE-genes between control and all patients with patients treated as biological replicates at FDR adjusted $P<0.05$ (ctrl vs. p1/p2/p14). When LBH-589-treated patients were compared to DMSO-treated controls (FDR<0.05), we considered previously identified DE-genes as "rescued" if they were not differentially expressed anymore. To assess the global effect of LBH-589, genes were identified as DE in DMSO vs. LBH-589 treatment, irrespective of the cell line, at FDR adjusted Padj<0.001. Network maps and pathways/cell components were generated in Ingenuity Pathway Analysis (IPA, Qiagen Redwood City, Version 01-07 (01-07)), using curated lists from Ingenuity Pathways Knowledge Base. Activation z-scores were used to infer the activation states ("increased" or "decreased") of implicated biological functions. Differentially expressed genes were annotated using the OMIM morbid map (version November 2017).

ChIP-seq analysis

MACS2 ChIP-seq peaks of mMSL2, mMOF and mKANSL3[21] were split in TSS (gene overlapping within ± 1kb of the TSS), gene body (gene overlapping from TSS+1kb until gene end) or distal (closest gene to any peak) groups based on the NCBIM37 Gencode annotation (v1) and compared to the DE-genes in *MSL3* patients. Mouse orthologues of *MSL3*-DE genes were identified on Ensembl BioMart. Overlaps, frequency calculations, statistical analyses and plots were generated in R (Version 1.1.383).

Scratch assay of HDFs

Equal number of cells were seeded in Culture insert-2 (IBIDI^{TM}) in 12 wells and let to attach. Cells were starved overnight before adding fresh media with LBH-589 (2 μM) or DMSO. The cell-free gaps in which the cell migration can be visualized was created by removal of the silicon comb. Pictures were taken with a conventional CCD camera on a regular inverted cell culture microscope with a 10x objective at 0, 6, 24 and 48h after removal of the comb and drug administration. Images were analysed using ImageJ (1.47v) measuring the cell-free gap and plotted as ratio over time 0 for each corresponding cell line.

H4K16ac Chromatin Immunoprecipitation (ChIP)

HDFs were starved overnight and collected by centrifugation after cell counting. Cells were crosslinked using 1% formaldehyde for 10 minutes. Sodium Butyrate was added to all buffers. Cell were fractionated (78840, Thermo Fisher) and nuclei dissolved in lysis buffer (20 mM Tris pH 8, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.25 M Sucrose, 1% Triton-X100, 10 μM EDTA, 0.5% SDS). 50'000 lysed nuclei were diluted with 5X dilution buffer (10 mM Tris-HCl pH8, 140 mM NaCl, 0.1% NaDeoxycholate, 1% Triton-X100, 1 mM EDTA, 1x protease, 0.1% SDS) and chromatin sheared using a Diagenode Bioruptor Pico (7 cycles with 30 sec ON/OFF. IP timing and antibody concentration were tightly set up using Flow cytometry (50'000 nuclei, 2 hours, RT, 1:750 dilution). DynaBeads-bound chromatin was washed using 2x IP dilution buffer, 1x High Salt wash buffer (400 mM NaCl), 1x LiCl buffer and 1x TE. Chromatin was reverse crosslinked at 65°C overnight followed by RNAseA and Proteinase K digestion. DNA was purified using phenol-chloroform extraction followed by Ethanol precipitation.

Statistics

All statistics were calculated with GraphPad Prism 6/7 or R (v1.1.383). One-way-ANOVA was performed for multiple comparisons followed by Bonferroni corrections for individual pairing. Dot plots are presented with center line as average±SEM. In the box plots in Figure 3c and Supplementary Figure 2e, the line that divides the box into 2 parts represents the median, the ends the upper and lower quartiles (first and third quartile). The extreme lines shows the highest and lowest value excluding

outliers. Bar plots represent the average±SEM with overlaid data points representing independent experiments. Results were considered significant at FDR (p-value) below 5%. ns: not significant. NA: not analyzed. SEM: standard error of the mean. Additional informations and test results of statistical analysis are provided in the Figure Legends and Supplementary Table 5.

Web Resources

NHLBI Exome Sequencing Project Exome Variant Server:
http://evs.gs.washington.edu/EVS/
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/
Picard: www.picard.sourceforge.net
University of Burgundy Centre de Calcul, https://haydn2005.u-bourgogne.fr/dsi-ccub/
Seattle Seq Annotation tool: snp.gs.washington.edu/SeattleSeqAnnotation137/
ExAC browser: http://exac.broadinstitute.org/
GATK, HaplotypeCaller:
https://www.broadinstitute.org/gatk/gatkdocs/org_broadinstitute_gatk_tools_walkers_haplotypecaller_HaplotypeCaller.php
EBI Expression Atlas: https://www.ebi.ac.uk/gxa/home/
Ensembl Biomart: https://www.ensembl.org/biomart/
Integrative Genome Viewer: https://software.broadinstitute.org/software/igv/

Reporting Summary

Further information on experimental design and reagents can be found in the Life Sciences Reporting Summary.

Data availability

RNA-seq data have been deposited to the Gene Expression Omnibus under the accession number GSE102250.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[72] with the dataset identifier PXD009317.

The damaging variants reported as disease-causing in this article were deposited in ClinVar database under study accession SUB2871008.

Methods-only References

53. Lelieveld, S. H. *et al.* Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat. Neurosci.* **19,** 1194–1196 (2016).
54. Samocha, K. E. *et al.* A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **46,** 944–950 (2014).
55. Gilissen, C. *et al.* Genome sequencing identifies major causes of severe intellectual disability. *Nature* **511,** 344–347 (2014).
56. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542,** 433–438 (2017).
57. Retterer, K. *et al.* Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* **18,** 696–704 (2016).
58. Thevenon, J. *et al.* Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin. Genet.* **89,** 700–707 (2016).
59. Pajusalu, S., Reimand, T. & Õunap, K. Novel homozygous mutation in KPTN gene causing a familial intellectual disability-macrocephaly syndrome. *Am. J. Med. Genet. A* **167,** 1913–1915 (2015).
60. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
61. Thåström, A. *et al.* Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **288,** 213–229 (1999).
62. Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375,** 23–44 (2004).
63. Basilicata, M. F., Frank, M., Solter, D., Brabletz, T. & Stemmler, M. P.

Inappropriate cadherin switching in the mouse epiblast compromises proper signaling between the epiblast and the extraembryonic ectoderm during gastrulation. *Sci. Rep.* **6,** 26562 (2016).

64. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1,** 2856–2860 (2006).
65. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2,** 1896–1906 (2007).
66. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13,** 731–740 (2016).
67. Kuhn, M. Contributions from Jed Wing and Steve Weston and Andre Williams and Chris Keefer and Allan Engelhardt and Tony Cooper and Zachary Mayer and Brenton Kenkel and the R Core Team and Michael Benesty and Reynald Lescarbeau and Andrew Ziem and Luca Scrucca., caret: Classification and Regression Training. in *R Package Version 6.0--47* (2015).
68. Stemmler, M. P., Hecht, A. & Kemler, R. E-cadherin intron 2 contains cis-regulatory elements essential for gene expression. *Development* **132,** 965–976 (2005).
69. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12,** 357–360 (2015).
70. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30,** 923–930 (2014).
71. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and

dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

72. Vaudel, M. *et al.* Exploring the potential of public proteomics data. *Proteomics* **16,** 214–225 (2016).

Figure 1

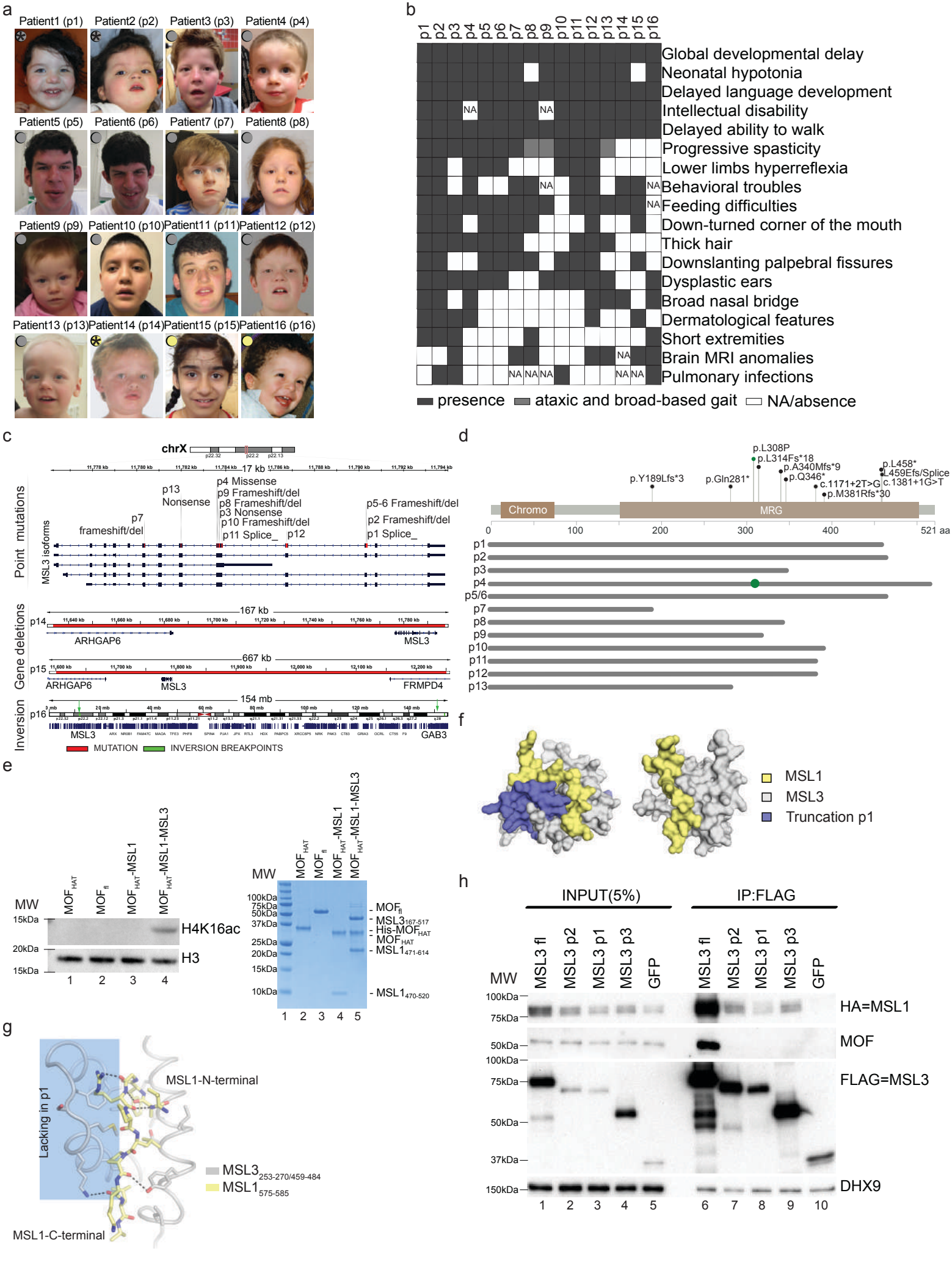

a
Patient1 (p1)
Patient2 (p2)
Patient3 (p3)
Patient4 (p4)
Patient5 (p5)
Patient6 (p6)
Patient7 (p7)
Patient8 (p8)
Patient9 (p9)
Patient10 (p10)
Patient11 (p11)
Patient12 (p12)
Patient13 (p13)
Patient14 (p14)
Patient15 (p15)
Patient16 (p16)
b
p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12 p13 p14 p15 p16
Global developmental delay
Neonatal hypotonia
Delayed language development
Intellectual disability
Delayed ability to walk
Progressive spasticity
Lower limbs hyperreflexia
Behavioral troubles
Feeding difficulties
Down-turned corner of the mouth
Thick hair
Downslanting palpebral fissures
Dysplastic ears
Broad nasal bridge
Dermatological features
Short extremities
Brain MRI anomalies
Pulmonary infections
NA
presence
ataxic and broad-based gait
NA/absence
c
chrX
17 kb
Point mutations
MSL3 isoforms
p13 Nonsense
p4 Missense
p9 Frameshift/del
p8 Frameshift/del
p3 Nonsense
p10 Frameshift/del
p11 Splice_
p12
p7 frameshift/del
p5-6 Frameshift/del
p2 Frameshift/del
p1 Splice_
Gene deletions
167 kb
p14
ARHGAP6
MSL3
667 kb
p15
ARHGAP6
MSL3
FRMPD4
Inversion
154 mb
p16
MSL3
GAB3
MUTATION
INVERSION BREAKPOINTS
d
p.L308P
p.L314Fs*18
p.A340Mfs*9
p.Q346*
p.M381Rfs*30
p.Y189Lfs*3
p.Gln281*
c.1171+2T>G
p.L458*
L459Efs/Splice
c.1381+1G>T
Chromo
MRG
0
100
200
300
400
521 aa
p1
p2
p3
p4
p5/6
p7
p8
p9
p10
p11
p12
p13
e
MW
MOF HAT
MOF fl
MOF HAT-MSL1
MOF HAT-MSL1-MSL3
15kDa
20kDa
15kDa
H4K16ac
H3
1 2 3 4
MW
100kDa
75kDa
50kDa
37kDa
25kDa
20kDa
15kDa
10kDa
- MOF fl
- MSL3 167-517
- His-MOF HAT
MOF HAT
- MSL1 471-614
- MSL1 470-520
1 2 3 4 5
f
MSL1
MSL3
Truncation p1
g
Lacking in p1
MSL1-N-terminal
MSL3 253-270/459-484
MSL1 575-585
MSL1-C-terminal
h
INPUT(5%)
IP:FLAG
MW
MSL3 fl
MSL3 p2
MSL3 p1
MSL3 p3
GFP
100kDa
75kDa
50kDa
100kDa
75kDa
50kDa
37kDa
150kDa
HA=MSL1
MOF
FLAG=MSL3
DHX9
1 2 3 4 5 6 7 8 9 10

Figure 2

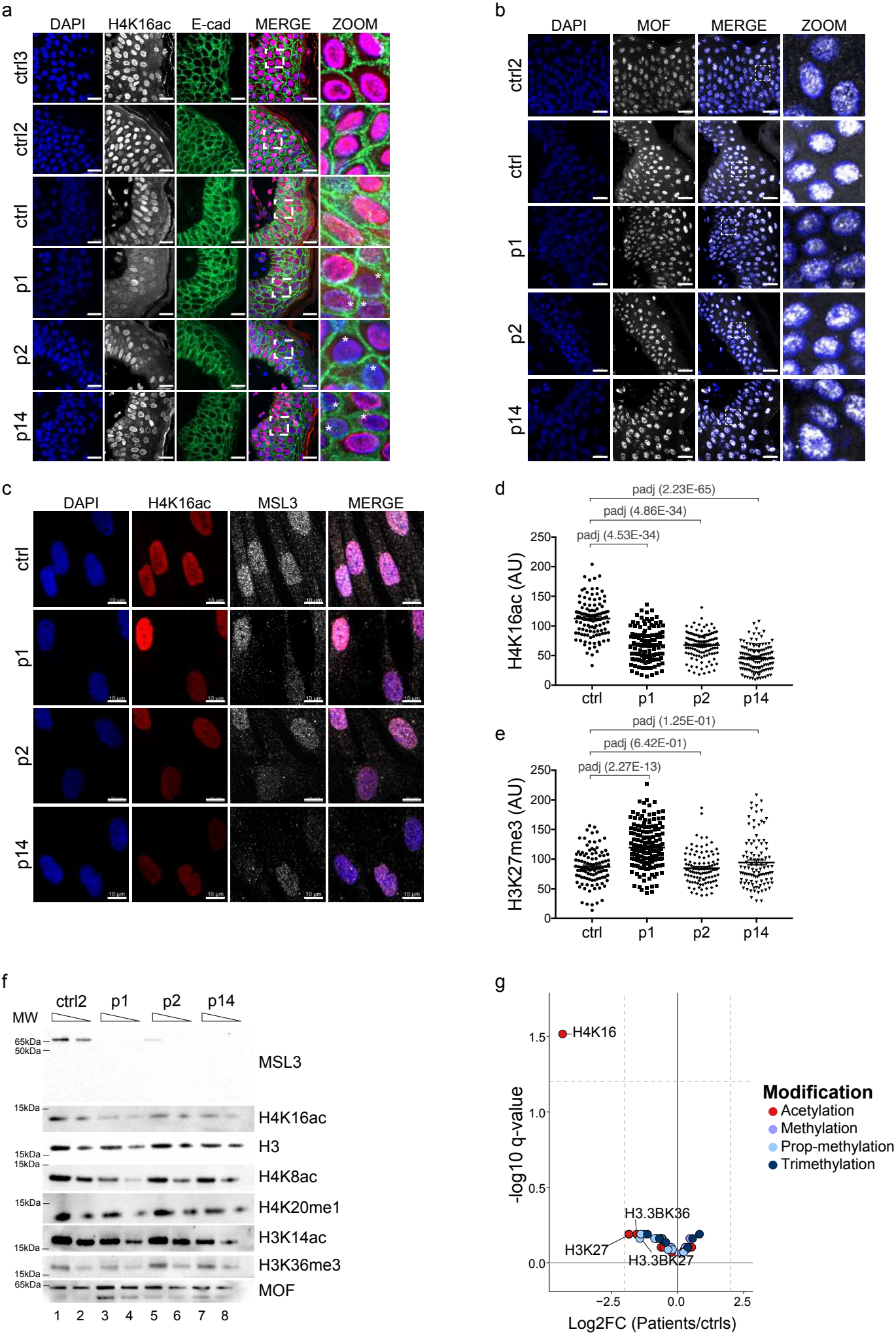

a
DAPI
H4K16ac
E-cad
MERGE
ZOOM
ctrl3
ctrl2
ctrl
p1
p2
p14
b
DAPI
MOF
MERGE
ZOOM
ctrl2
ctrl
p1
p2
p14
c
DAPI
H4K16ac
MSL3
MERGE
ctrl
p1
p2
p14
d
padj (2.23E-65)
padj (4.86E-34)
padj (4.53E-34)
H4K16ac (AU)
ctrl
p1
p2
p14
e
padj (1.25E-01)
padj (6.42E-01)
padj (2.27E-13)
H3K27me3 (AU)
ctrl
p1
p2
p14
f
MW
ctrl2
p1
p2
p14
65kDa
50kDa
MSL3
15kDa
H4K16ac
H3
H4K8ac
H4K20me1
H3K14ac
H3K36me3
MOF
1 2 3 4 5 6 7 8
g
H4K16
H3.3BK36
H3K27
H3.3BK27
-log10 q-value
Log2FC (Patients/ctrls)
Modification
Acetylation
Methylation
Prop-methylation
Trimethylation

Figure 3

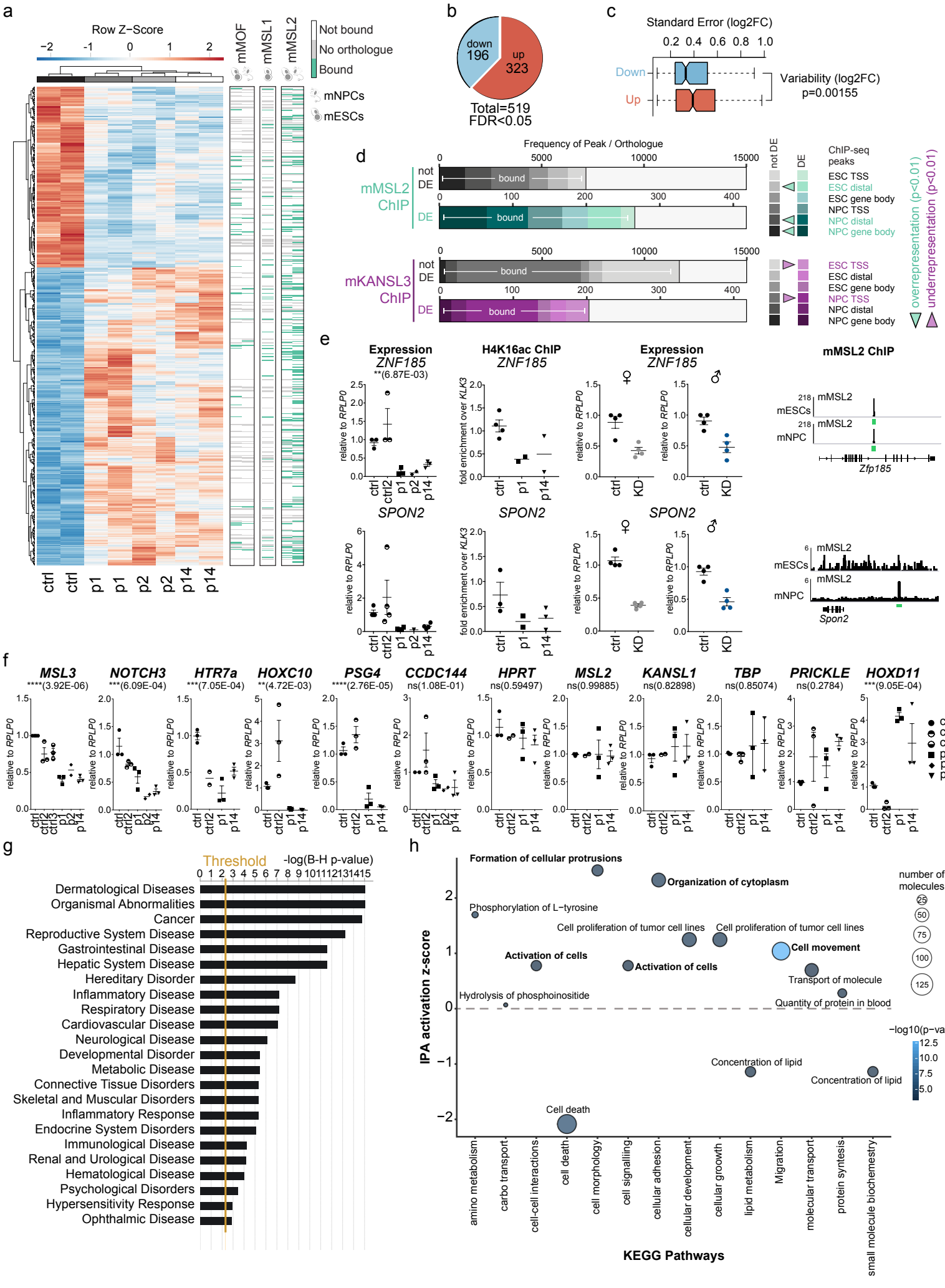
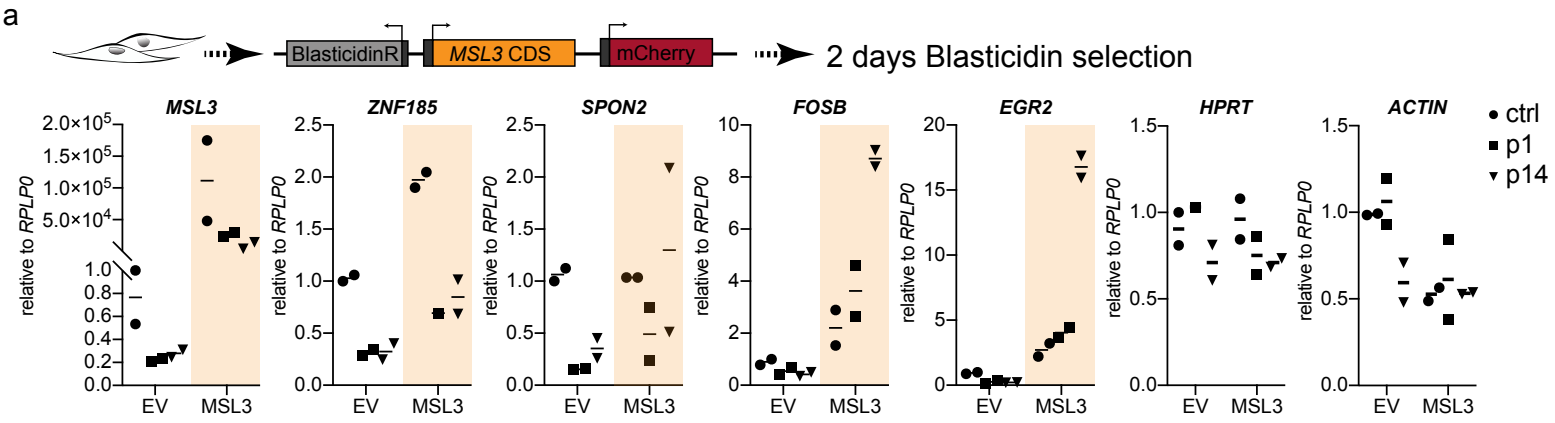
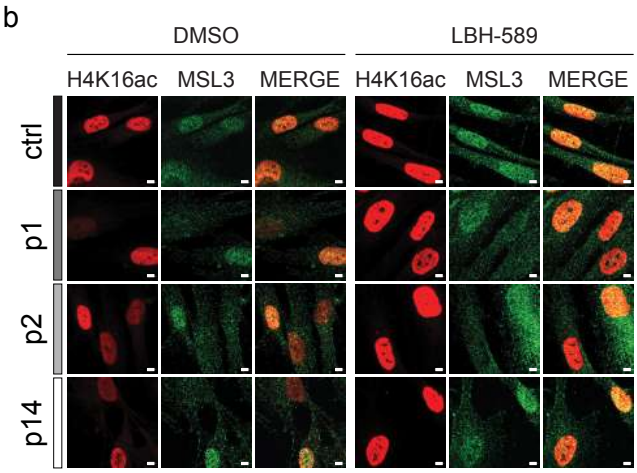
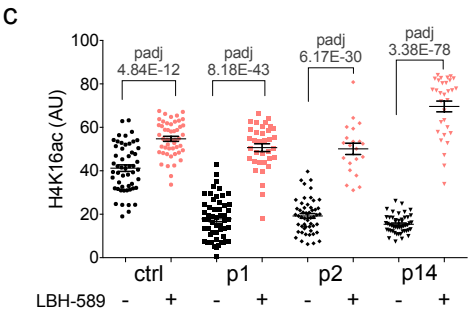
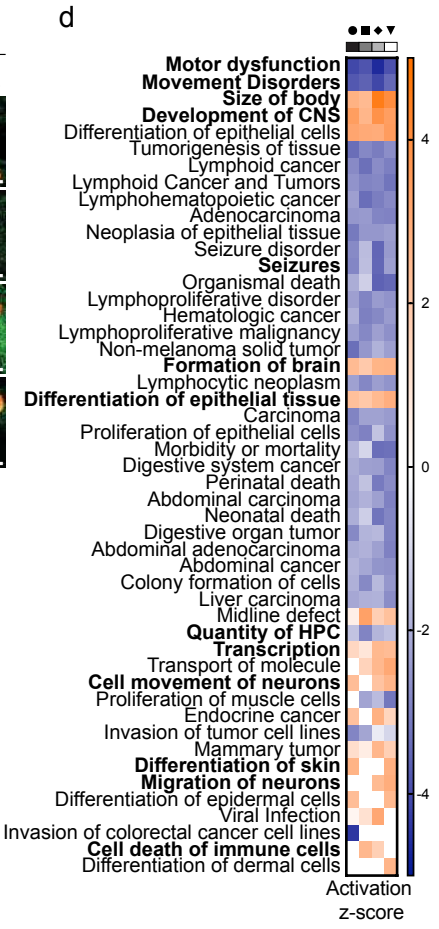

a
Row Z−Score
−2 -1 0 1 2
mMOF
mMSL1
mMSL2
Not bound
No orthologue
Bound
mNPCs
mESCs
ctrl ctrl p1 p1 p2 p2 p14 p14
b
down 196
up 323
Total=519
FDR<0.05
c
Standard Error (log2FC)
0.2 0.4 0.6 0.8 1.0
Down
Up
Variability (log2FC)
p=0.00155
d
Frequency of Peak / Orthologue
mMSL2 ChIP
mKANSL3 ChIP
not DE
DE
bound
ChIP-seq peaks
ESC TSS
ESC distal
ESC gene body
NPC TSS
NPC distal
NPC gene body
overrepresentation (p<0.01)
underrepresentation (p<0.01)
e
Expression
ZNF185
**(6.87E-03)
H4K16ac ChIP
ZNF185
Expression
ZNF185
mMSL2 ChIP
SPON2
SPON2
SPON2
Zfp185
Spon2
mESCs
mNPC
mMSL2
f
MSL3 ****(3.92E-06)
NOTCH3 ***(6.09E-04)
HTR7a ***(7.05E-04)
HOXC10 **(4.72E-03)
PSG4 ****(2.76E-05)
CCDC144 ns(1.08E-01)
HPRT ns(0.59497)
MSL2 ns(0.99885)
KANSL1 ns(0.82898)
TBP ns(0.85074)
PRICKLE ns(0.2784)
HOXD11 ***(9.05E-04)
relative to RPLP0
ctrl ctrl2 ctrl3 p1 p2 p14
g
Threshold
-log(B-H p-value)
Dermatological Diseases
Organismal Abnormalities
Cancer
Reproductive System Disease
Gastrointestinal Disease
Hepatic System Disease
Hereditary Disorder
Inflammatory Disease
Respiratory Disease
Cardiovascular Disease
Neurological Disease
Developmental Disorder
Metabolic Disease
Connective Tissue Disorders
Skeletal and Muscular Disorders
Inflammatory Response
Endocrine System Disorders
Immunological Disease
Renal and Urological Disease
Hematological Disease
Psychological Disorders
Hypersensitivity Response
Ophthalmic Disease
h
Formation of cellular protrusions
Organization of cytoplasm
Phosphorylation of L−tyrosine
Cell proliferation of tumor cell lines
Cell proliferation of tumor cell lines
Cell movement
Activation of cells
Activation of cells
Transport of molecule
Hydrolysis of phosphoinositide
Quantity of protein in blood
Concentration of lipid
Concentration of lipid
Cell death
number of molecules
−log10(p−value)
IPA activation z-score
amino metabolism
carbo transport
cell-cell interactions
cell death
cell morphology
cell signalling
cellular adhesion
cellular development
cellular growth
lipid metabolism
Migration
molecular transport
protein syntesis
small molecule biochemestry
KEGG Pathways

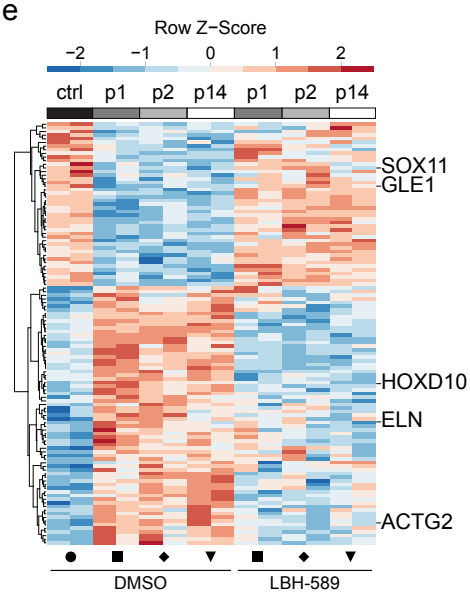
Figure 4

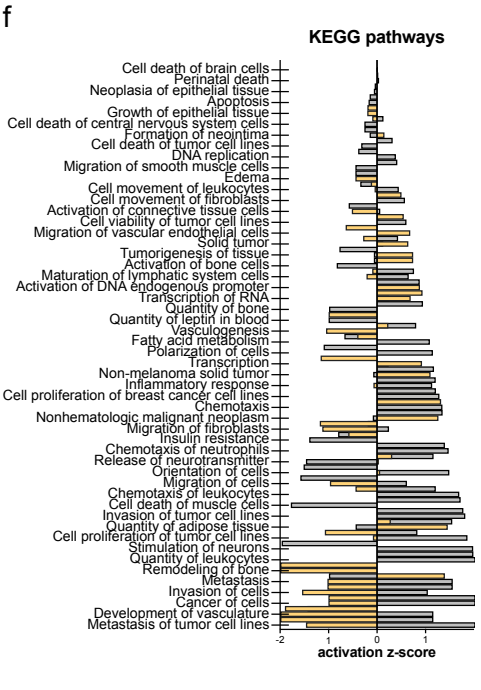

a
BlasticidinR
MSL3 CDS
mCherry
2 days Blasticidin selection
MSL3
ZNF185
SPON2
FOSB
EGR2
HPRT
ACTIN
relative to RPLP0
EV
MSL3
ctrl
p1
p14

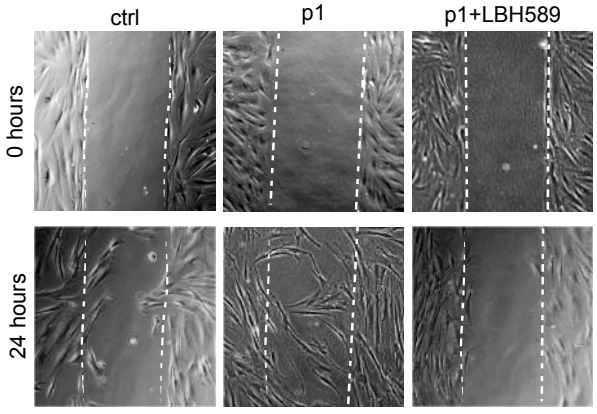

b
DMSO
LBH-589
H4K16ac
MSL3
MERGE
ctrl
p1
p2
p14

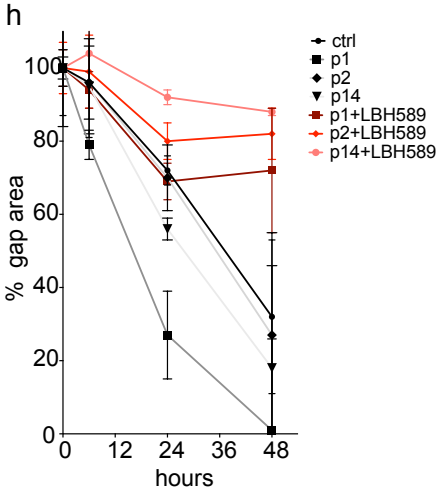

c
H4K16ac (AU)
padj 4.84E-12
padj 8.18E-43
padj 6.17E-30
padj 3.38E-78
ctrl
p1
p2
p14
LBH-589

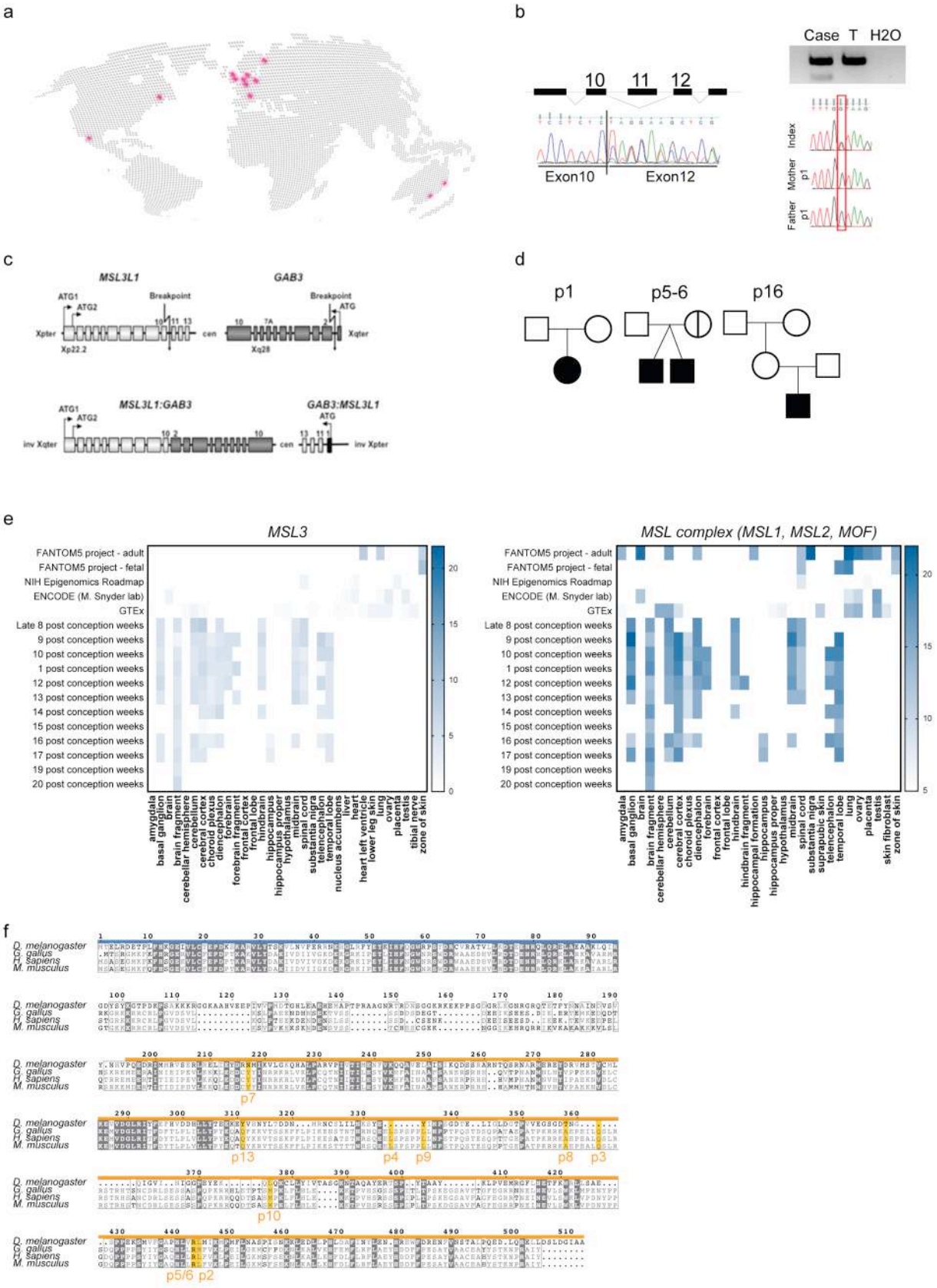

d
Motor dysfunction
Movement Disorders
Size of body
Development of CNS
Differentiation of epithelial cells
Tumorigenesis of tissue
Lymphoid cancer
Lymphoid Cancer and Tumors
Lymphohematopoietic cancer
Adenocarcinoma
Neoplasia of epithelial tissue
Seizure disorder
Seizures
Organismal death
Lymphoproliferative disorder
Hematologic cancer
Lymphoproliferative malignancy
Non-melanoma solid tumor
Formation of brain
Lymphocytic neoplasm
Differentiation of epithelial tissue
Carcinoma
Proliferation of epithelial cells
Morbidity or mortality
Digestive system cancer
Perinatal death
Abdominal carcinoma
Neonatal death
Digestive organ tumor
Abdominal adenocarcinoma
Abdominal cancer
Colony formation of cells
Liver carcinoma
Midline defect
Quantity of HPC
Transcription
Transport of molecule
Cell movement of neurons
Proliferation of muscle cells
Endocrine cancer
Invasion of tumor cell lines
Mammary tumor
Differentiation of skin
Migration of neurons
Differentiation of epidermal cells
Viral Infection
Invasion of colorectal cancer cell lines
Cell death of immune cells
Differentiation of dermal cells
Activation z-score


e
Row Z-Score
ctrl
p1
p2
p14
SOX11
GLE1
HOXD10
ELN
ACTG2
DMSO
LBH-589


f
KEGG pathways
Cell death of brain cells
Perinatal death
Neoplasia of epithelial tissue
Apoptosis
Growth of epithelial tissue
Cell death of central nervous system cells
Formation of neointima
Cell death of tumor cell lines
DNA replication
Migration of smooth muscle cells
Edema
Cell movement of leukocytes
Cell movement of fibroblasts
Activation of connective tissue cells
Cell viability of tumor cell lines
Migration of vascular endothelial cells
Solid tumor
Tumorigenesis of tissue
Activation of bone cells
Maturation of lymphatic system cells
Activation of DNA endogenous promoter
Transcription of RNA
Quantity of bone
Quantity of leptin in blood
Vasculogenesis
Fatty acid metabolism
Polarization of cells
Transcription
Non-melanoma solid tumor
Inflammatory response
Cell proliferation of breast cancer cell lines
Chemotaxis
Nonhematologic malignant neoplasm
Migration of fibroblasts
Insulin resistance
Chemotaxis of neutrophils
Release of neurotransmitter
Orientation of cells
Migration of cells
Chemotaxis of leukocytes
Cell death of muscle cells
Invasion of tumor cell lines
Quantity of adipose tissue
Cell proliferation of tumor cell lines
Stimulation of neurons
Quantity of leukocytes
Remodeling of bone
Metastasis
Invasion of cells
Cancer of cells
Development of vasculature
Metastasis of tumor cell lines
activation z-score
LBH-589 Treatment
Deregulated in Patients


g
ctrl
p1
p1+LBH589
0 hours
24 hours

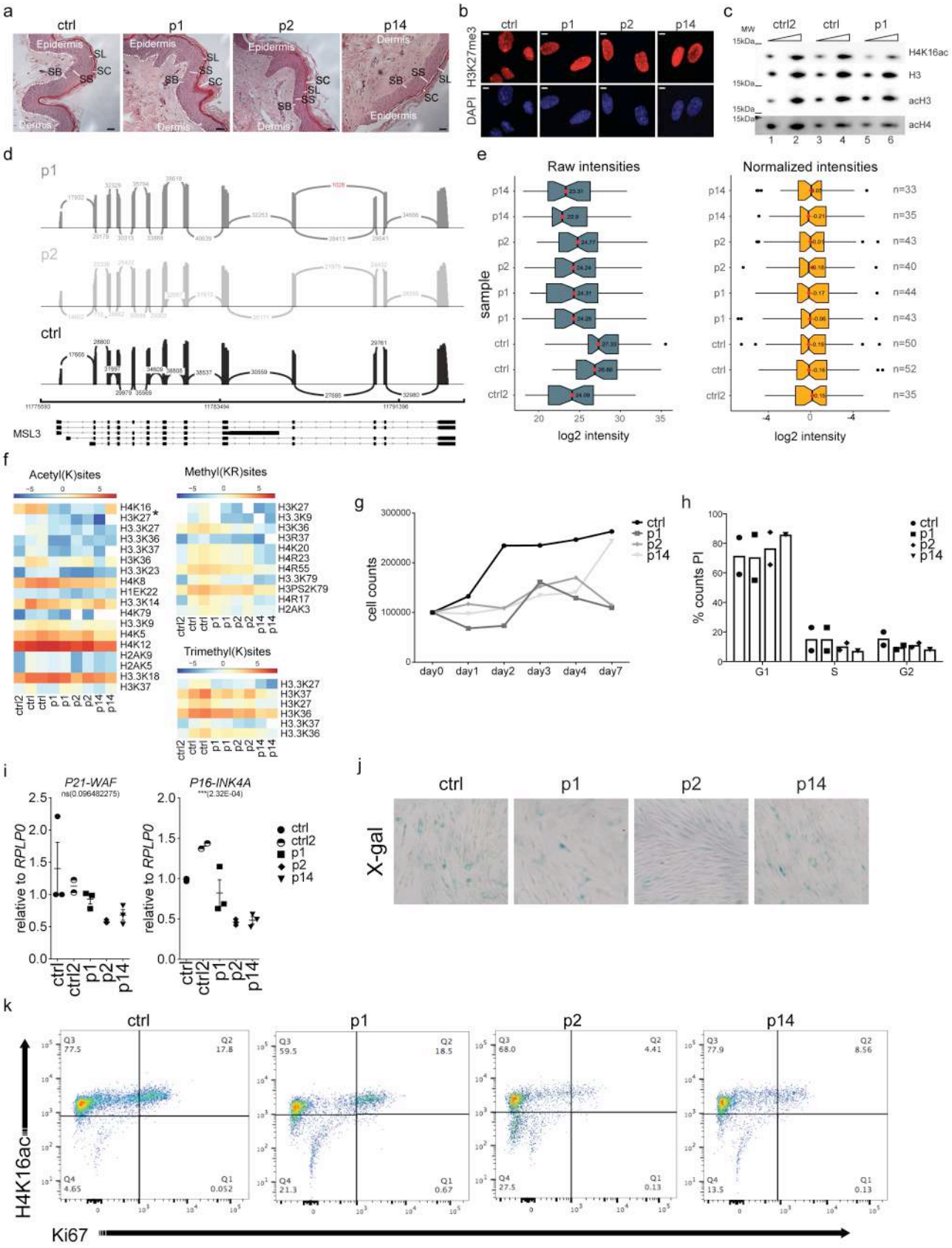

h
% gap area
hours
ctrl
p1
p2
p14
p1+LBH589
p2+LBH589
p14+LBH589


a
b
Case T H2O
10 11 12
Exon10 Exon12
Index
Mother p1
Father p1
c
MSL3L1
GAB3
ATG1
ATG2
Breakpoint
Xpter
cen
Xqter
Xp22.2
Xq28
MSL3L1:GAB3
GAB3:MSL3L1
inv Xqter
d
p1
p5-6
p16
e
MSL3
MSL complex (MSL1, MSL2, MOF)
FANTOM5 project - adult
FANTOM5 project - fetal
NIH Epigenomics Roadmap
ENCODE (M. Snyder lab)
GTEx
Late 8 post conception weeks
9 post conception weeks
10 post conception weeks
1 post conception weeks
12 post conception weeks
13 post conception weeks
14 post conception weeks
15 post conception weeks
16 post conception weeks
17 post conception weeks
19 post conception weeks
20 post conception weeks
amygdala
basal ganglion
brain fragment
cerebellar hemisphere
cerebellum
cerebral cortex
choroid plexus
diencephalon
forebrain
forebrain fragment
frontal cortex
frontal lobe
hindbrain
hippocampus
hippocampus proper
hypothalamus
midbrain
spinal cord
substantia nigra
telencephalon
temporal lobe
nucleus accumbens
liver
heart
heart left ventricle
lower leg skin
lung
ovary
placenta
testis
tibial nerve
zone of skin
hindbrain fragment
hippocampal formation
suprapubic skin
skin fibroblast
f
D. melanogaster
G. gallus
H. sapiens
M. musculus
p7
p13
p4
p9
p8
p3
p10
p5/6
p2

Supplementary Figure 1

MSL3 variants cause a novel syndrome

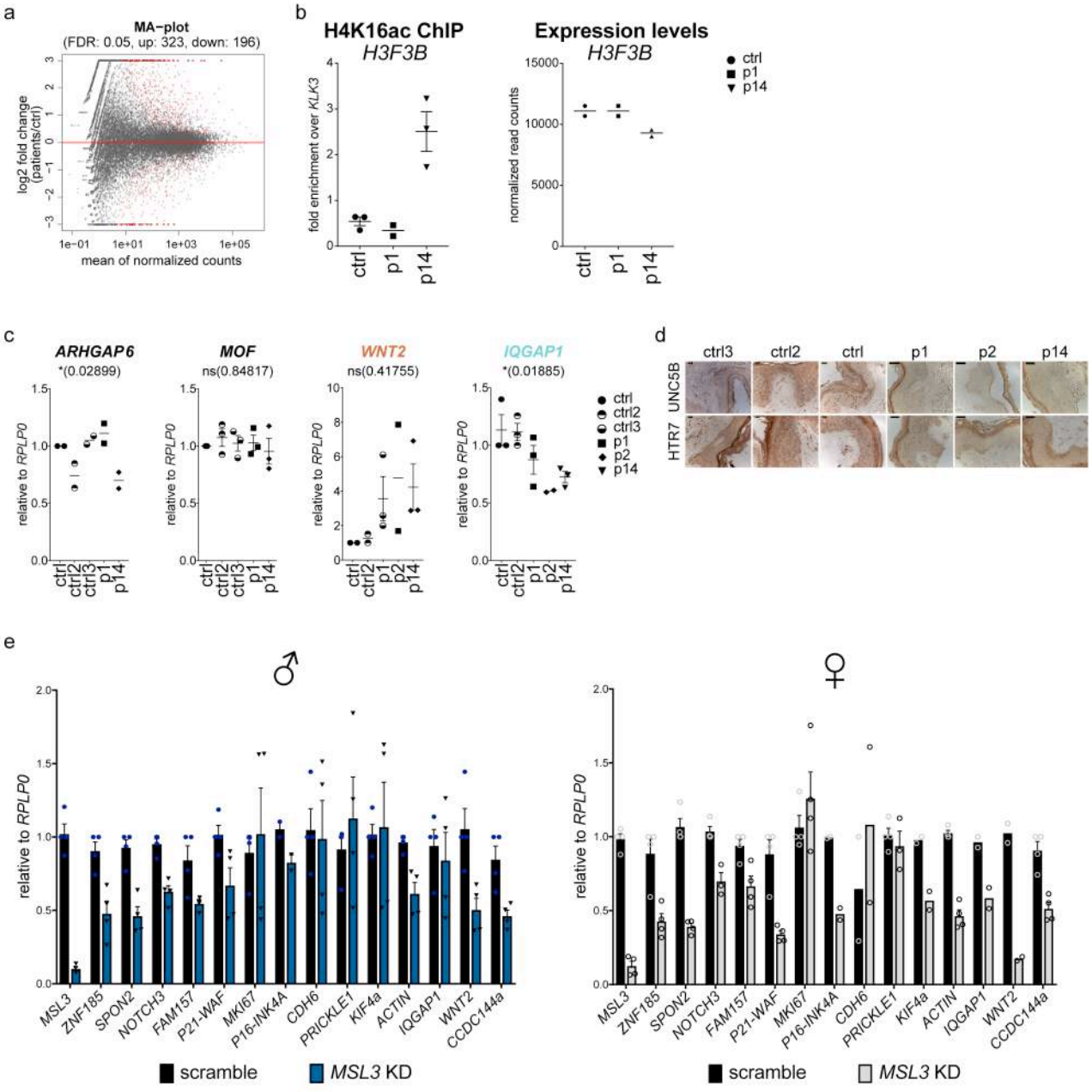
(a) Representation of the *MSL3* patient countries of origin. (b) Sanger sequencing confirming the skipping of *MSL3* exon 11 in p1 cDNA. The allelic ratio was in favour of random X inactivation. Right: Cropped agarose gel picture of RT-PCR products. (c) Schematic representation of X-chromosome inversion of p16. (d) Patients pedigrees. No *MSL3* variant could be detected in the parents of patients p1-11 (the father of patient p10 could not be tested). p16's mother is reported as mildly affected with no clinical details provided. Her X chromosome inversion occurred *de novo*. (e) Heatmap representing RNA-seq data for *MSL3* in the developing human brain compared to other members of the MSL complex. Data was retrieved from https://www.ebi.ac.uk/gxa/home/ (f) Amino acid sequence alignment of MSL3 orthologues showing that the mutated residues are highly conserved throughout evolution.


a
ctrl
p1
p2
p14
Epidermis
Dermis
SL
SC
SB
SS
b
H3K27me3
DAPI
c
MW
ctrl2
ctrl
p1
15kDa
H4K16ac
H3
acH3
acH4
1 2 3 4 5 6
d
p1
p2
ctrl
MSL3
e
Raw intensities
Normalized intensities
sample
log2 intensity
f
Acetyl(K)sites
Methyl(KR)sites
Trimethyl(K)sites
H4K16
H3K27
H3.3K27
H3.3K36
H3.3K37
H3K36
H3.3K23
H4K8
H1EK22
H3.3K14
H4K79
H3.3K9
H4K5
H4K12
H2AK9
H2AK5
H3.3K18
H3K37
H3K27
H3.3K9
H3K36
H3R37
H4K20
H4R23
H4R55
H3.3K79
H3PS2K79
H4R17
H2AK3
H3.3K27
H3K37
H3K27
H3K36
H3.3K37
H3.3K36
g
cell counts
day0
day1
day2
day3
day4
day7
h
% counts PI
G1
S
G2
i
P21-WAF
ns(0.096482275)
P16-INK4A
***(2.32E-04)
relative to RPLP0
ctrl
ctrl2
p1
p2
p14
j
X-gal
k
H4K16ac
Ki67

Supplementary Figure 2

Characterization of *MSL3* patient-derived fibroblasts

(a) FFPE-skin sections from ctrl and p1/p2/p14 stained with H&E. Dashed line distinguishes dermis and epidermis layers. Architectural skin layers are demarcated. Scale bar=20 µm. Patients donated *n*=1 skin sample, at least two sections per slide were analysed. SB=Stratum Basale, SL= Stratum lucidum, SS=Stratum Spinosum, SC=Stratum Corneum (b) Immunostaining for H3K27me3 (red) in primary dermal fibroblasts (HDFs). Scale bar=5 µm. The staining was repeated twice with similar results. (c) Cropped immunoblot for H4K16ac, total Histone H3, as well as H3 and H4 pan-acetylation in additional HDF lines. The experiment was repeated twice with similar results. (d) Sashimi plot derived from MiSeq results showing exon skipping in p1, but not p2 or ctrl HDF cDNA. (e) Distribution of identified protein intensities measured in LC-MS/MS experiments before (left) and after normalization (right). Protein intensities as well as modified site intensities were normalized and scaled by adjusting the centers of the distributions around zero to account for loading differences in SDS-PAGE. Box plots are centered in the median with the lower and upper hinges corresponding to the first and third quartiles. Normalized values were used for further statistical analysis as described in Supplementary Table 2. (f) Heatmap representing all acetyl (K), mono and tri-methyl (R-K) histone modifications normalized intensities detected over bulk histone background level as in Supplementary Table 2. (g) Proliferation curve in p1, p2 and p14 compared to ctrl. The center value at each time point represents the average of *n*=2 independent experiments. (h) FACS cell cycle analysis of ctrl and p1/p2/p14 HDFs, Propidium Iodide was used to define cell cycle phases. Bar plots represents average of *n*=2 independent experiments with overlaid data points. (i) RT-qPCR analysis of senescence markers *P16-INK4A* and *P21-WAF* displayed as dot plots. Expression levels were normalized to *RPLP0* and expressed relative to ctrl. Each data point represents an independent experiment (*n*) with center line representing average±SEM when applicable. p-values were determined by Ordinary One-way ANOVA followed by Bonferroni multiple corrections. Further details and statistical test values are provided in Supplementary Table 5. (j) Representative DIC images of β-galactosidase activity assay performed in ctrl and p1/p2/p14 HDFs. The experiment was repeated 3 times with similar results. (k) Representative FACS analysis of MKI67 (x-axis) and H4K16ac (y-axis) in ctrl and p1/p2/p14 HDFs. Quadrants show % of cells with relative abundance of cell populations. The experiment was repeated twice with similar results.

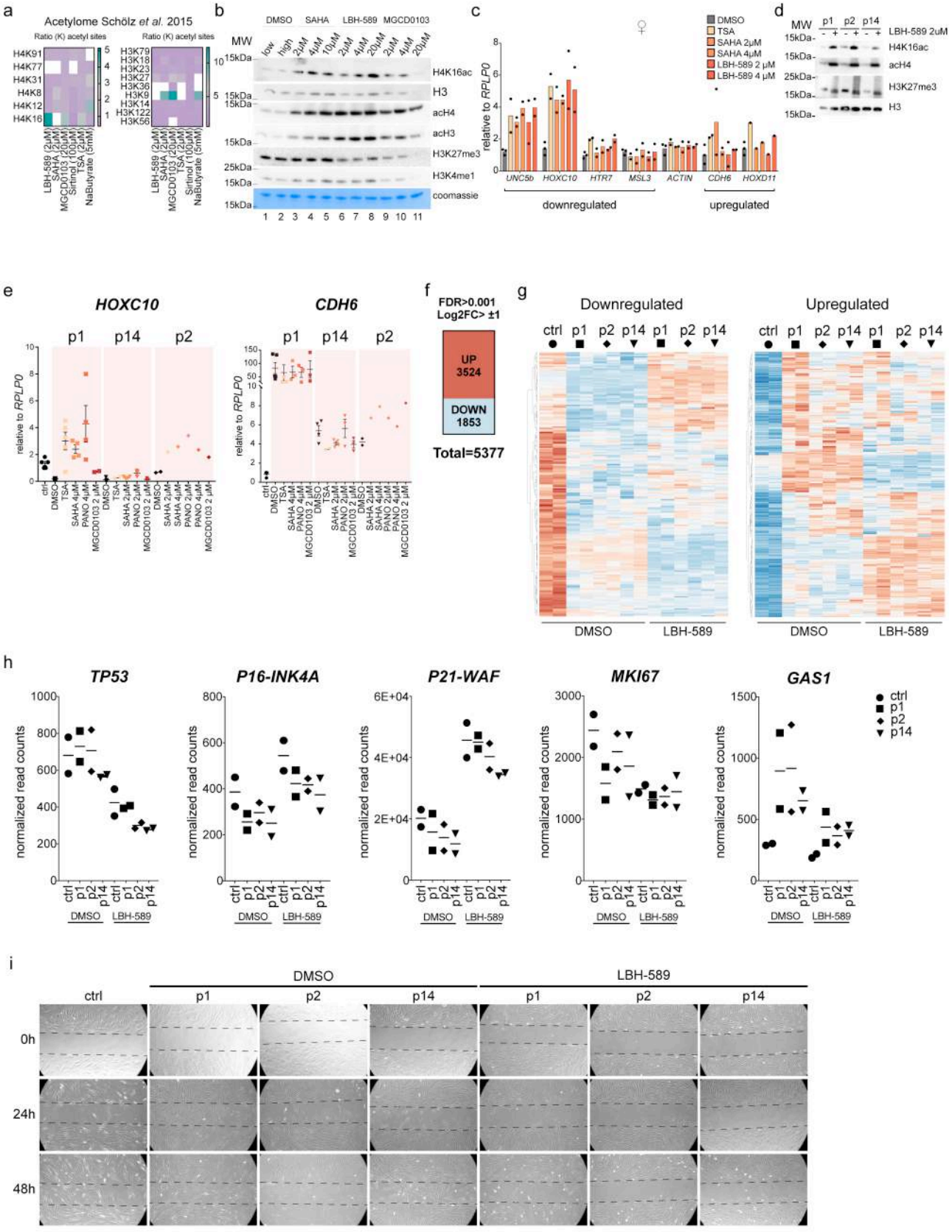


Supplementary Figure 3

Validation of transcriptional responses in *MSL3* patients

(a) MA-plot comparing the mean of the normalized counts versus the log2FC obtained from RNA-seq of patients versus ctrl HDFs (*n*=2 passages of ctrl were compared with *n*=2 passages of each p1, p2 and p14). DE-genes (FDR cutoff of 0.05) are marked in red. (b) H4K16ac ChIP qPCR analysis of *H3F3B* and respective expression levels from RNA-seq (normalized read counts) displayed as dot plots. H4K16ac ChIP-qPCR enrichment values were calculated relative to input and expressed as a fold change enrichment over the negative control *KLK3*. Each data point represents an independent experiment (*n*) with center line representing average±SEM where applicable. (c) RT-qPCR expression analysis in HDFs displayed as dotplots. Expression levels were normalized to *RPLP0* and expressed relative to ctrl. Each data point represents an independent experiment (*n*) with center line representing average±SEM where applicable. p-values were determined by Ordinary One-way ANOVA followed by Bonferroni multiple corrections. Further details and statistical test values are provided in Supplementary Table 5. (d) Immunohistochemistry for Serotonin receptor (HTR7) and Netrin

receptor (UNC5B) on control and patient-derived FFPE-skin sections. Similar staining results were obtained in *n*=2 sections per slide (e) RT-qPCR expression analysis in male and female HDFs upon *MSL3* knock-down (KD) displayed as bar plots representing the average±SEM. Expression levels were normalized to *RPLP0* and expressed relative to scrambled siRNA. The same data points for *ZNF185* and *SPON2* are also shown in main Figure 3 and are illustrated again for comparative purposes. Each overlaid data point represents the number (*n*) of independent experiments.


a
Acetylome Schölz et al. 2015
Ratio (K) acetyl sites
H4K91
H4K77
H4K31
H4K8
H4K12
H4K16
H3K79
H3K18
H3K23
H3K27
H3K36
H3K9
H3K14
H3K122
H3K56
LBH-589 (2uM)
SAHA (2uM)
MGCD0103 (2uM)
Sirtinol (100uM)
TSA (2uM)
NaButyrate (5mM)
b
DMSO
SAHA
LBH-589
MGCD0103
MW
low
high
2µM
4µM
10µM
2µM
4µM
20µM
2µM
4µM
20µM
15kDa
25kDa
H4K16ac
H3
acH4
acH3
H3K27me3
H3K4me1
coomassie
1 2 3 4 5 6 7 8 9 10 11
c
♀
DMSO
TSA
SAHA 2µM
SAHA 4µM
LBH-589 2 µM
LBH-589 4 µM
relative to RPLP0
UINC5b
HOXC10
HTR7
MSL3
ACTIN
CDH6
HOXD11
downregulated
upregulated
d
MW
p1
p2
p14
LBH-589 2uM
15kDa
25kDa
H4K16ac
acH4
H3K27me3
H3
e
HOXC10
CDH6
p1
p14
p2
relative to RPLP0
ctrl
DMSO
TSA
SAHA 4µM
PANO 4µM
MGCD0103 2 µM
SAHA 2µM
PANO 2µM
f
FDR>0.001
Log2FC> ±1
UP
3524
DOWN
1853
Total=5377
g
Downregulated
Upregulated
ctrl
p1
p2
p14
DMSO
LBH-589
h
TP53
P16-INK4A
P21-WAF
MKI67
GAS1
normalized read counts
ctrl
p1
p2
p14
DMSO
LBH-589
i
DMSO
LBH-589
ctrl
p1
p2
p14
0h
24h
48h

Supplementary Figure 4

Response to HDACi in *MSL3* patient cells

(a) Heatmap representing histone modification changes upon HDACi treatments (dataset from Schölz et al., Nat. Biotechnol. 2015, doi: 10.1038/nbt.3130). (b) Cropped immunoblot for H4K16ac, pan-acetylated H3 and H4, H3K27me3, H3K4me1 and H3 of nuclear extracts of ctrl HDFs treated with HDACi. The same extracts were separated on a coomassie-stained gel serving as loading control. The experiment was repeated twice with similar results. (c) RT-qPCR of acetylation-sensitive targets in ctrl HDF. The bar plot represents the average of *n*=2 independent experiments with overlaid individual data points. (d) Cropped immunoblot of p1/p2/p14 HDFs nuclear extracts upon LBH-589 treatment. The experiment was repeated twice with similar results. (e) RT-qPCR of p1/p2/p14 HDFs treated with four different HDACi: SAHA (vorinostat, HDAC class I and II inhibitor), LBH-589 (panobinostat, pan-HDACi) and MGCD0103 (mocetinostat, HDACi class I and IV). Expression levels are normalized to *RPLP0* and calculated relative to ctrl (without treatment) and shown as dot plots. Each data point represents (*n*) independent experiments with center line representing the average±SEM where applicable. (f) Scheme representing the number of DE-genes upon LBH-589 treatment. (g) Heatmap representing z-scores on the *MSL3* patients DE-downregulated (left, *n*=196) and DE-upregulated (right, *n*=323) genes upon LBH-589 obtained by RNA-seq (p1/p2/p14, 2 passages). (h) Dot plots of normalized RNA-seq read counts for DNA damage and cell cycle marker genes before and after treatment with LBH-589. The center line represents the average of *n*=2 independent experiments. (i) Representative DIC images upon LBH-589 treatment of ctrl and p1/p2/p14 HDFs at 0h, 24h, 48h after creating a gap area.

Figure 1e

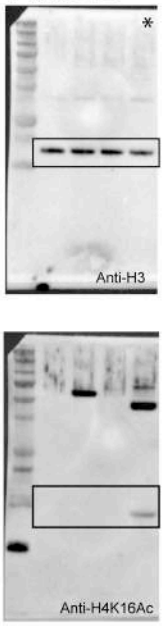


Figure 1h

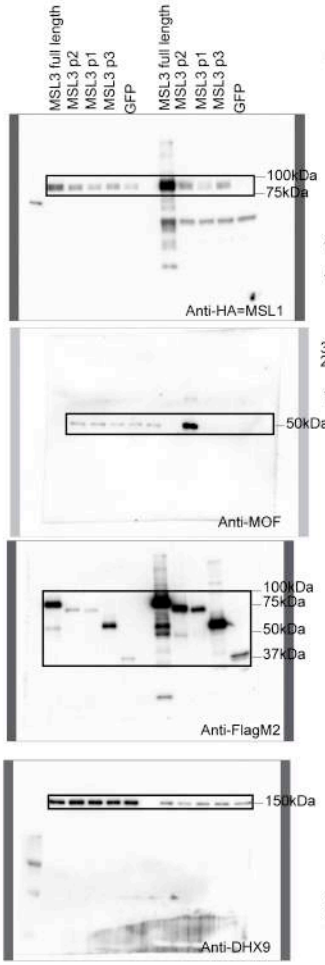
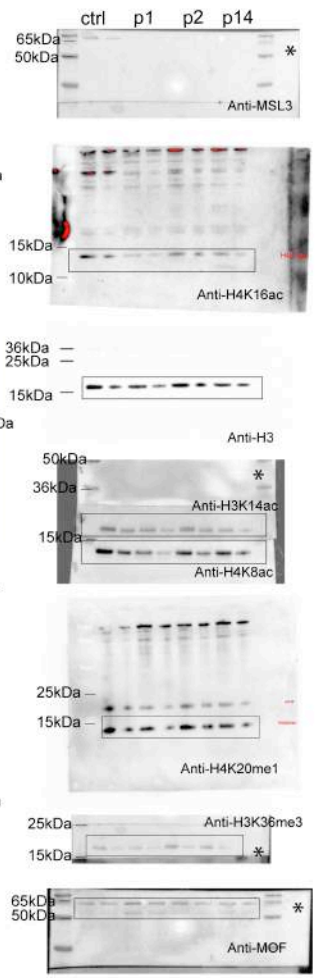
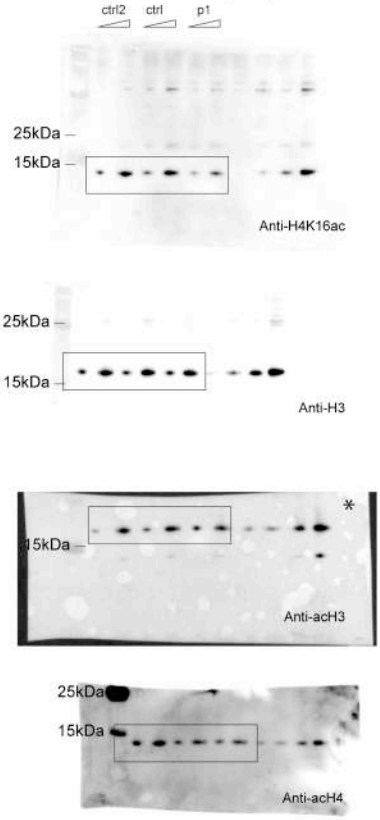


Figure 2f

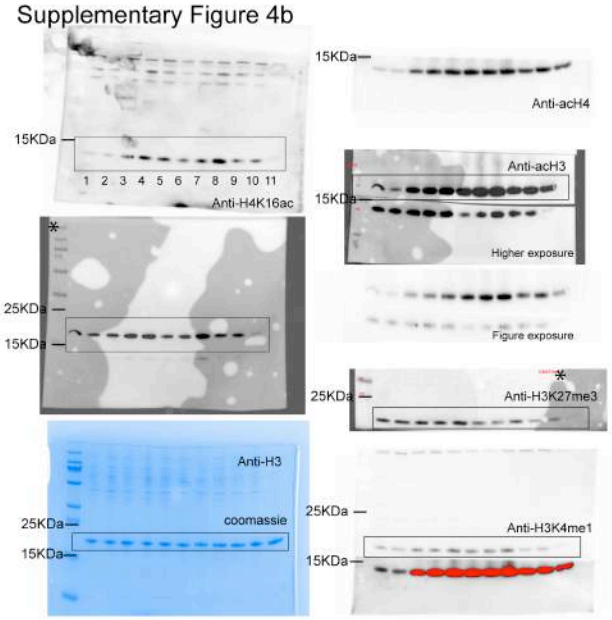


Supplementary Figure 2c

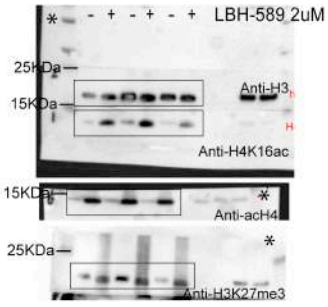


Supplementary Figure 1b

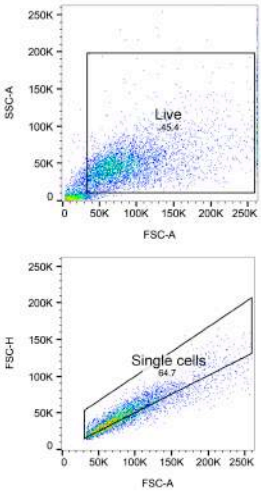
Supplementary Figure 4b



Supplementary Figure 4d



Supplementary Figure 2 h and k



Supplementary Figure 5

Uncropped western blots, agarose and coomassie gel pictures, and gating strategy example.

To illustrate molecular weight markers, epi-white and chemiluminescence pictures were merged in the display (respective blots are marked with an asterisk). Actual figure panels represent only chemiluminescence and not the merged pictures. Cropped regions are framed. Representative gating strategy for excluding debris (upper), and doublets (lower) in flow cytometry analysis of 1×104 events. Singlets are defined on the Forward-Scatter (height) vs. Forward-Scatter (area) dot plot.